

DOCUMENT RESUME

ED 262 737

HE 018 824

AUTHOR Novalis, Carol
TITLE Validation of Student and Parent Report Data on the Basic Grant Application Form. Final Report. Volume II, Estimated Income: CRT Look-Up Study, Individual Case Follow-Up Study, IRS Sample Study.
INSTITUTION Applied Management Sciences, Inc., Silver Spring, Md.
SPONS AGENCY Bureau of Student Financial Assistance (DHEW/OE), Washington, D.C.
PUB DATE 7 Jan 77
CONTRACT 300-75-0227
NOTE 190p.; For related documents, see HE 018 812-826.
PUB TYPE Reports - Evaluative/Feasibility (142) -- Statistical Data (110)

EDRS PRICE MF01/PC08 Plus Postage.
DESCRIPTORS Accountability; *College Students; Disclosure; *Eligibility; *Error Patterns; *Financial Aid Applicants; Financial Needs; Grants; Higher Education; *Income; Program Evaluation; Student Characteristics; *Student Financial Aid; Student Records

IDENTIFIERS *Basic Educational Opportunity Grants

ABSTRACT

The use of estimated income to analyze financial need of applicants to the Basic Educational Opportunity Grant (BEOG) program was investigated. Attention was focused on: how well applicants estimate their income; reasons for errors in estimation, and whether applicants supplying income tax returns supply true versions. For 1,547 eligible BEOG applicants, estimates of income on a 1974-1975 supplemental application were compared were estimates made on regular 1975-1976 applications. Findings included: those with lower incomes were more accurate in estimating income, while those with higher incomes more often overestimated; the later in the year the supplemental application was filed, the more accurate the income estimate; accuracy was not related to the date in which the 1975 application was filed; and as the rate of accurate estimation increased, the size and the frequency of discrepancies became smaller. Compared to the first study, a followup study of individual cases found 11 percent fewer applicants who estimated accurately within \$300. An additional study compared 1974-1975 estimated income with 1974 income tax return data for 2,840 eligible applicants and 1,750 ineligible. Another analysis assessed the accuracy of estimation for applicants judged to be ineligible based on their supplemental BEOG applications. (SW)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

HE

ED 262737

G-59

VALIDATION OF STUDENT AND PARENT
REPORT DATA ON THE BASIC
GRANT APPLICATION FORM

FINAL REPORT

VOLUME II - ESTIMATED INCOME STUDIES

CRT Look-up Study
Individual Case Follow-up Study
IRS Sample Study

January 7, 1977

Prepared for:

Division of Basic and State Student Grants
Office of Education
Department of Health, Education, and Welfare

Under Contract No. 500-75-0227

AF 018824

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

- ☒ This document has been reproduced as received from the person or organization originating it.
- ☐ Minor changes have been made to improve reproduction quality.
- Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.



Author: Carol Novalis, M.A.

Project Director: Martin Bloom, M.A.

Principal Investigator: JoAnn Kuchak, M.A.

Director of Data Processing: Michael Consiglio

ACKNOWLEDGEMENTS

Applied Management Sciences wishes to express its sincere gratitude to Ernst Becker, Project Officer, and to Sylvia Cowan of the Office of Education, for their input and review of all project plans and deliverables, and to the following people:

Office of Education

Peter K.U. Voigt
Donna Whinston
Steven Jenks

Internal Revenue Service

Paul Grayson
Gloria Koteen
Jack Francis
Thelma Adams

The American College Testing Program

Shannon Janes
Judy Hoeger
Randy Sidwell
Dennis Scott
Gwen Burnett

Applied Management Sciences

Todd Tucker	Jean Harman
Martin Bloom	Andrea Epstein
JoAnn Kuchak	Cynthia Maus
Ronald Vogel	Laura Webb
Michael Consiglio	Beatrice Bray
Karen Smith	Colleen Crump
Carol Novalis	June Hamning
John Tesch	Ellen Shea
Richard Weiskopf	Donna Nicholson
Robert Goodman	Deborah Sullivan
Jeanne Adler	Janice McNeil
Mike Puma	Carolyn Cutchin
Carole Schor	Ada Rippberger

TABLE OF CONTENTS

<u>Chapter</u>	<u>Title</u>	<u>Page</u>
1	INTRODUCTION.	1.1
2	THE CRT LOOK-UP STUDY	2.1
	2.1: STUDY OBJECTIVES AND ANALYSIS QUESTIONS. . . .	2.1
	2.2: CRT LOOK-UP METHODOLOGY.	2.2
	2.2.1: CRT Look-up Sampling Procedures. . . .	2.2
	2.2.2: CRT Look-Up Methodology.	2.3
	2.3: CRT LOOK-UP STUDY FINDINGS.	2.3
	2.3.1: Characteristics of the Sample.	2.4
	2.3.2: Findings Related to the Analysis Questions.	2.5
	2.4: CRT LOOK-UP SUMMARY AND CONCLUSIONS	2.21
3	INDIVIDUAL CASE FOLLOW-UP	3.1
	3.1: STUDY OBJECTIVES AND ANALYSIS QUESTIONS . . .	3.2
	3.2: INDIVIDUAL CASE FOLLOW-UP METHODOLOGY	3.3
	3.2.1: Individual Case Follow-up Sampling Procedures	3.3
	3.2.2: Individual Case Follow-up Methodology.	3.4
	3.3: INDIVIDUAL CASE FOLLOW-UP FINDINGS.	3.6
	3.3.1: Characteristics of the Sample.	3.6
	3.3.2: Findings Related to the Analysis Questions.	3.9
	3.4: INDIVIDUAL CASE FOLLOW-UP CONCLUSIONS	3.20
4	IRS SAMPLE STUDY.	4.1
	4.1: IRS SAMPLE STUDY OBJECTIVES AND ANALYSIS QUESTIONS	4.1
	4.2: IRS SAMPLE STUDY METHODOLOGY.	4.3
	4.2.1: IRS Sample Study Sample Selection. . .	4.3
	4.2.2: Confidentiality Measures	4.5
	4.2.3: Analysis Methodology	4.8
	4.2.4: Specification of Variables	4.10
	4.3: IRS SAMPLE STUDY FINDINGS	4.12

TABLE OF CONTENTS (Continued)

<u>Chapter</u>	<u>Title</u>	<u>Page</u>
	4.3.1: Characteristics of the Population. . . .	4.14
	4.3.2: Findings Related to the Analysis Questions.	4.17
	4.3.2.1: Overall Findings Obtained From IRS Sample Study.	4.17
	4.3.2.2: Findings Obtained for the Eligible Study Group.	4.26
	4.3.2.3: Findings Obtained for the Ineligible Study Group.	4.71
	4.4: SUMMARY AND CONCLUSIONS	4.102
5	OVERALL FINDINGS OF THE ESTIMATED INCOME STUDIES. . .	5.1

LIST OF EXHIBITS

<u>Exhibits</u>	<u>Title</u>	<u>Page</u>
1.1	1974-1975 Basic Education Opportunity Grant Program Supplemental Application.	1.4
2.1	Distribution of Misreporting Found in CRT-Look-Up Study	2.7
2.2	Frequency of AGI Discrepancies - CRT Look-Up Study	2.8
2.3	Differences in 1974 Income By Date of Filing 1974-75 Supplemental Application.	2.11
2.4	Frequency of AGI Discrepancies By Date of Filing 1974 Supplemental Application.	2.12
2.5	Frequencies of Discrepancies in 1974 Income By Date of Filing 1975-76 Application	2.15
2.6	Frequency of AGI Discrepancies By Date of Filing 1975-76 BEOG Application.	2.16
2.7	AGI Discrepancies By Adjusted Gross Income Level. .	2.18
2.8	Frequency of AGI Discrepancies By Adjusted Gross Income Level.	2.19
3.1	Frequency of AGI Discrepancies - Individual Case Follow-Up Study	3.11
3.2	Reason for Filing Supplemental By Size of Discrepancy in 1974 Income	3.12
3.3	Frequency of AGI Discrepancies By Supplemental Reason.	3.13
3.4	Frequency of Discrepancies Between Estimate and IRS Data Provided By Applicant.	3.16
3.5	Frequency of AGI Discrepancies By Final AGI	3.17
4.1	Population and Sample Sizes (Estimated Income). . .	4.6
4.2	IRS Study - Characteristics of the Population and Study Group	4.15

TABLE OF EXHIBITS (Continued)

<u>Exhibit</u>	<u>Title</u>	<u>Page</u>
4.3	Overall Accuracy of Estimated Income Population. . . .	4.19
4.4	Frequency of Discrepancies by Supplemental Reason - Total Study Group.	4.22
4.5	Overall Accuracy of Student Eligibility Indices Based on AGI Plus Taxes Paid Substitutions -- Total Study Group.	4.24
4.6	Frequency of Discrepancies in Student Eligibility Indices - Total Study Group.	4.25
4.7	Overall Accuracy -- Eligible Study Group	4.27
4.8	Distribution of Income Levels for IRS -- Ineligible Portions of the BEOG Study Group	4.29
4.9	Discrepancies in Adjusted Gross Income for the Eligible Study Group	4.33
4.10	Frequency of Adjusted Gross Income Discrepancies By BEOG-Reported AGI - Eligible Dependents.	4.35
4.11	Frequency of Adjusted Gross Income Discrepancies By BEOG-Reported AGI - Eligible Dependents.	4.36
4.12	Frequency of Adjusted Gross Income Discrepancies By Supplemental Reason - Eligible Study Group	4.38
4.13	Frequency of Discrepancies in Adjusted Gross Income By the Date of Supplemental Signatures of Applicants - Eligibles.	4.41
4.14	Frequency of Discrepancies in Taxes Paid By Dependency Status - Eligibles.	4.44
4.15	Distribution of Discrepancies In Student Eligibility Index For Each Substitution Analysis	4.58
4.16	Distribution of Student Eligibility Index Discrepancies By Dependency Status and Type of IRS-Substitution - Eligible Study Group	4.60
4.17	Distribution of Student Eligibility Index Discrepancies By Level of Adjusted Gross Income - Eligible Dependents	4.61

LIST OF EXHIBITS (Continued)

<u>Exhibit</u>	<u>Title</u>	<u>Page</u>
4.18	Distribution of Student Eligibility Index Discrepancies By Level of Adjusted Gross Income - Eligible Independents.	4.63
4.19	Frequency of Discrepancies in Student Eligibility Indices Found in AGI Substitution Analysis By BEOG-Reported Eligibility Index - Eligibles.	4.64
4.20	Frequency of Discrepancies in Student Eligibility Indices Found in the AGI and Taxes Paid Substitution Analysis By BEOG-Reported SEI - Eligibles	4.65
4.21	BEOG-Reported Eligible and IRS-Reported Ineligible Applications With Estimated Income By BEOG-Reported AGI and Dependency Status	4.69
4.22	Overall Accuracy of Adjusted Gross Income Estimates - Ineligibles Study Group	4.72
4.23	Frequency of Adjusted Gross Income Discrepancies By Dependency Status - Ineligibles	4.74
4.24	Frequency of Adjusted Gross Income Discrepancies By Level of AGI -- Independent Ineligibles	4.76
4.25	Frequency of Discrepancies in Adjusted Gross Income By Level of AGI -- Dependent Ineligibles	4.77
4.26	Frequency of Discrepancies in Adjusted Gross Income By Supplemental Reason - Ineligibles.	4.80
4.27	Frequency of Discrepancies in Adjusted Gross Income By Date of Supplemental Signature	4.82
4.28	Frequency of Discrepancies in Taxes Paid By Dependency Status - Ineligibles	4.85
4.29	Summary of Overall Discrepancy Rates in Student Eligibility Indices Obtained In IRS-Substitution Analyses.	4.90
4.30	Frequency of Discrepancies in Student Eligibility Index By Dependency Status - Ineligibles.	4.94

LIST OF EXHIBITS (Continued)

<u>Exhibit</u>	<u>Title</u>	<u>Page</u>
4.31	Frequency of Discrepancies In Student Eligibility Indices By Adjusted Gross Income - Dependent Ineligibles.	4.95
4.32	Frequency of Discrepancies In Student Eligibility Indices By Adjusted Gross Income - Independent Ineligibles.	4.96
4.33	Frequency of Discrepancies In Student Eligibility Indices By SEI Discrepancy - Ineligibles	4.98
4.34	Number and Percent of BEOG-Reported Ineligible and IRS-Reported Eligible Applicants With Estimated Income By BEOG-Reported AGI and Dependency Status. .	4.101

1

INTRODUCTION

The Basic Grant Validation Study was funded in May 1975 to investigate the extent, types and causes of misreporting on applications for Basic Educational Opportunity Grants. Basic Grants were established by Congress for the purpose of assisting postsecondary students in obtaining the benefits of postsecondary education. Thus, the Grant is designed to promote the equality of educational opportunity among all students by bolstering the ability of the most needy to pay for their education. Misreporting, therefore, would tend to cause funds to be diverted from those students who have the greatest need. By determining the causes of misreporting, the Office of Education hopes to be able to reduce its frequency by making changes in the reporting procedures and the applications.

This Validation Study has examined the types of misreporting such as errors in transcription, misunderstanding of application instructions, and wilful errors. This project has had several components, each taking a different approach to locating discrepancies. The present volume describes the component study which investigated the reporting accuracy of a small segment of the applicant population which was allowed to estimate income for the year which the Grant is awarded, rather than reporting the income of the previous year.

The issue of using estimated versus actual income for financial need analysis purposes has long been a topic of debate in the financial

aid community. In the private sector the two major financial need analysis services, Education Testing Service and the American College Testing Program, collect both an estimate of the current year's income, and a report of the actual income entered on the most recent tax form filed. The estimate is necessary since both the ETS and ACT services begin processing before the previous year's tax form is filed. These services include an edit routine in their processing which enables an analysis of unusual or unreasonable changes in income. The Basic Grant program, however, begins application processing in mid-February and March when many applicants and their families have already completed a Federal Income Tax Return and after all income from the previous year is earned. Therefore, the Basic Grant program is able to use the actual income received during the previous year as an indicator of financial strength. Further, since the actual prior year's income is reported, it is possible to validate the amount reported.

The Basic Grant policy of using actual instead of estimated income was derived from empirical evidence. Several studies, such as Boyd and Himelick^{1/} have shown the prior year's income to be the best practical indicator of the income resources available to the student in the current academic year. These studies have pointed out that, when requested to provide an income estimate, low income families tend to overestimate and middle income families tend to underestimate. In both cases the effect on aid allocation is in the undesired direction.

The Basic Grant Program does, however, recognize that in particular circumstances the prior year's income is not the best

^{1/} Boyd, Joseph D. and Himelick, Byron L. "A Five-Part Paper Dealing with the use of Actual or Estimated Income in Calculating Family Contribution;" prepared under contract OEC-0-74-8610 for the Division of Basic Grants, Office of Education, 1974.

See also Orwig & Jones, "Can Financial Need Analysis be Simplified?", The American College Testing Program, Iowa City, Iowa.

predictor of current financial strength. Specifically, Basic Grant applicants are allowed to file a supplemental application and estimate their current income when specific types of catastrophic events have occurred and have severely affected the present year's income, causing an increased need for financial aid. Strict criteria are used by Basic Grants in determining who is permitted to submit a supplemental form for income estimation in a given year. Exhibit 1.1 displays the application form utilized by estimated income applicants. The BEOG applications for the 1974-75 Grant year utilized the following criteria:

- A parent or spouse whose 1973 income was included in the calculation of the student eligibility index died during 1973 or 1974; or
- Such a parent or spouse has experienced a loss of employment of at least ten consecutive weeks in 1973-74 or at least 10 total weeks in 1974; 1/ or
- Such a parent, spouse, or applicant has been unable to pursue normal income producing activities due to disabilities or natural disaster for at least ten consecutive weeks in 1974; or
- Since the time of application the parents of the applicant have become separated or divorced; or
- Since the time of application, the applicant has become separated or divorced.

As part of the validation study of the Basic Grant applications, it was important to assess the accuracy of this estimation process. The Estimated Income portion of the study has the following objectives:

- To determine how well people can estimate.
- To provide empirical evidence as to reasons for errors in estimation.
- To determine if individuals supplying income tax returns as part of the documentation required in the individual case follow-up studies supply true versions of the IRS 1040.
- To recommend ways to improve the estimation system.

1/ NOTE: The 1975-76 application was changed so that independent students who worked full time and became unemployed could estimate their income.

DE USE ONLY 	DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE OFFICE OF EDUCATION WASHINGTON, D.C. 20202	FORM APPROVED OMB NO. 11-00553
---	---	-----------------------------------

SUPPLEMENTAL FORM FOR 1974-75 ACADEMIC YEAR

(Basic Educational Opportunity Grant Program)

IMPORTANT: An Application for Determination of Basic Grant Eligibility must be submitted with this form.

PURPOSE: The purpose of this form is to allow for special recomputation of the Basic Grant Eligibility Index for those applicants whose financial circumstances have changed significantly since 1973.

CONDITIONS FOR FILING: To be eligible for such recomputation, the applicant must meet one of the following five conditions:

- a parent or spouse whose 1973 income was included in the calculation of the student eligibility index died during 1973 or 1974, OR
- such a parent or spouse has experienced a loss of employment of at least ten consecutive weeks in 1973-74 or at least ten total weeks in 1974, OR
- such a parent, spouse or applicant has been unable to

pursue normal income producing activities due to disability or natural disaster for at least ten consecutive weeks in 1974, OR

- since the time of application, the parents of the applicant have become separated or divorced, OR
- since the time of application, the applicant has become separated or divorced.

DIRECTIONS: You must submit an Application for Determination of Basic Grant Eligibility (for the 1974-75 year) with this Supplemental Form, even if you have already submitted one. If you have not already submitted one, you may submit your original application with this form. Special directions for completing the Application for Determination of Basic Grant Eligibility to send with this form are provided below. **REMEMBER:** Use 1974 information.

1. Applicant's name (last name first)

2. Social Security Number

3. Date of Birth

4. House Number and Street Address City State ZIP Code

5. HAVE YOU PREVIOUSLY FILED AN APPLICATION FOR DETERMINATION OF BASIC GRANT ELIGIBILITY FOR THE 1974-75 ACADEMIC YEAR? (Check One) ☐ YES ☐ NO

SECTION A: CONDITION FOR FILING A SUPPLEMENTAL FORM - COMPLETE THE BOX (A, B, C, OR D) THAT DESCRIBES YOUR REASON FOR SUBMITTING THIS FORM.

A. DATE OF PARENT OR SPOUSE'S DEATH
 Mo. Day Year

B. DATES OF LOSS OF EMPLOYMENT BY PARENT OR SPOUSE
 (Does Not Apply To Applicant)
 From Mo. Day Year To Mo. Day Year

C. DATES OF INABILITY OF APPLICANT, PARENT, OR SPOUSE TO PURSUE NORMAL INCOME PRODUCING ACTIVITIES DUE TO DISABILITY OR NATURAL DISASTER
 From Mo. Day Year To Mo. Day Year

D. DATE PARENTS WERE SEPARATED OR DIVORCED
 Mo. Day Year

E. DATE APPLICANT WAS SEPARATED OR DIVORCED
 Mo. Day Year

IF ITEM (A, B, OR C) WAS COMPLETED ABOVE, ENTER NAME OF APPROPRIATE PERSON
 (Parent, Spouse, or Applicant)
 (last name first)

1973 INCOME OF THAT PERSON
 \$.00

I (WE) CERTIFY that this Supplemental Form is accurate and complete to the best of my (our) knowledge. I (WE) agree to provide, if requested, documentation necessary to verify information reported on this form.

Signature of Applicant

Signature of Parent, Guardian or Spouse

Date

SPECIAL DIRECTIONS FOR COMPLETING APPLICATION FOR DETERMINATION OF BASIC GRANT ELIGIBILITY (DE FORM 255): COMPLETE APPLICATION ACCORDING TO THE INSTRUCTIONS INCLUDED IN THE APPLICATION PACKET, BUT WITH THE FOLLOWING EXCEPTIONS.

EXCEPTION 1 Item 17 from 19-20 Item 21 from 22-23	EXCEPTION 2 Item 22 from 23-26 Item 26 from 27-28	- use 1974 information, rather than 1973. - the amount of income should be the sum of (1) actual income from January 1, 1974 to the date of this application, and (2) estimated income from the date of this application through December 31, 1974. - Note: This item (and Federal Income Tax) will be completed by the processing agent. - the amount of income should be the sum of (1) actual income from January 1, 1974 to the date of this application, and (2) estimated income from the date of this application through December 31, 1974.
--	--	--

WARNING: ANY PERSON WHO KNOWINGLY MAKES A FALSE STATEMENT OR MISREPRESENTATION ON THIS FORM SHALL BE SUBJECT TO A FINE, OR TO IMPRISONMENT, OR TO BOTH, UNDER PROVISIONS OF THE UNITED STATES CRIMINAL CODE.

Send Application and Supplemental Form To:
 Basic Educational Opportunity Grant Program
 P.O. Box 2254
 Washington, D.C. 20013

DE FORM 255-2, 4/74 PREVIOUS EDITION IS OBSOLETE.

Because estimates are based largely on unknown factors, it is likely that a comparatively large proportion of the Estimated Income population will have substantial discrepancies on their applications. To prevent the number of discrepancies from being excessively large, the instructions for estimating income request that the applicant first report the total income received for that portion of the year preceding the date of filing the supplemental, then estimate the income expected during the remainder of the year, and finally, to add the two sums together. However, there has been no method suggested for estimating the future income, so applicants must rely on their own estimates. Herein lies the likelihood of substantial discrepancies from "true income." In order to determine the accuracy of the estimated income population, the following three component studies were designed and performed:

- The Internal Revenue Service Sample component was comprised of computerized comparisons of the data estimated on the Basic Grant application with actual data reported to IRS. The procedures for the estimated income study were part of the larger IRS study reported in Volume III.
- The Individual Follow-up Component procedures paralleled those of the Pre-established Criteria Study (Volume I). Individuals who estimated their income for the 1974-75 BEOG year were randomly selected and contacted by the Validation Contractor. These individuals were asked to supply a copy of the 1974 income tax return to support the application upon which their Grant award was based.
- The CRT Look-up Component investigated a sample of estimators on the 1974-75 application who also applied for a Grant in 1975-76. This was done by sampling recipients and subsequently determining via the cathode-ray terminal if a record existed for those individuals in the 1975-76 BEOG data base. If a record existed, the actual income earned in 1974 (reported on the 1975-76 application) was compared with the estimated 1974 income reported on the 1974-75 supplemental application.

The remainder of this volume presents the findings of the Estimated Income Study. Chapter 2 will describe the CRT Look-up

Study. The Individual Follow-up Study is presented in Chapter 3. Chapter 4 discusses the methodology and findings of the IRS components. Chapter 5 presents the conclusions to be drawn from the investigation of estimating accuracy and discusses recommendations for the Basic Grant estimating procedures resulting from the study findings.

2

THE CRT LOOK-UP STUDY

The purpose of the CRT Look-up Study was to compare the applications of eligible BEOG applicants in 1974-75 who submitted a supplemental form and estimated their 1974 income with the application of these individuals in 1975-76. Since the previous year's adjusted gross income is reported on the regular Basic Grant application, the regular 1975-76 applications contained the actual 1974 income. This was compared with the estimated 1974 income reported on the 1974-75 supplemental application to determine the accuracy of the estimated. Because a catastrophic event such as death of a significant contributor to the applicant's family income is not likely to occur two years in a row (although unemployment might) it was hypothesized that 1974-75 estimators would not file a supplemental form in 1975-76.

2.1: STUDY OBJECTIVES AND ANALYSIS QUESTIONS

The CRT Look-up provides information concerning the degree of accuracy of the estimation procedures, the subsequent behavior of estimators, and whether or not the date of filing either application is a variable related to the accuracy of reporting.

The CRT Look-up component of the Estimated Income Study provides insight into each of the main study objectives listed in Chapter 1. It provides an indication, based on a large sample, of the accuracy of the estimate compared with the income reported for the same year after that year is completed. The frequency and extent of errors by the categories of income and

filing dates of the applicants shows the relationship of these variables to accuracy. Based on the findings of this and other study components, recommendations for changes to the Estimated Income application procedures may be made.

The comparison between the two years' report of 1974 income is useful when related to other Validation studies in assessing overall discrepancy rates and in determining the accuracy of 1975-76 application reports of income. More specifically the following analysis questions are answered by the CRT Look-up study:

- 1) What was the overall accuracy of estimation?
- 2) Do individuals who file a supplemental form later in the Grant application year have a more accurate estimate of their income than those who file earlier?
- 3) How will the dates of filing the 1975-76 applications be distributed for applicants who had filed a supplemental form in the previous year?
- 4) How is the income level of the applicant related to the accuracy of estimation?

The next section of this chapter describes the sampling and analysis methodology for the CRT Look-up study. Remaining portions of the chapter will describe the findings and conclusions.

2.2: CRT LOOK-UP METHODOLOGY

The CRT Look-up study utilized a random sample of estimators from the 1974-75 BEOG master file of eligible applicants. A roster of these applicants was then compared to the 1975-76 BEOG master file to determine if these applicants had applied for a Basic Grant in 1975-76.

2.2.1: CRT Look-up Sampling Procedures

A sample of 1,547 individuals who filed a supplemental application and estimated their 1974 income was selected from all eligible BEOG applicants in the 1974-75 application year. This sample size represents an absolute error of less than 5 percent at the .95 level of confidence. The sample was drawn randomly by computer, utilizing a skip interval of seven with a random start.

Analyses were performed on that subset of the sample which met the following criteria:

- the applicant filed a regular application in 1975-76
- the applicant's financial dependency status was the same for both years.

A total of 962 applicants met these criteria.

2.2.2: CRT Look-Up Methodology

The 1975-76 master file was searched to locate each individual in the sample. This search utilized a cathode-ray tube (CRT) terminal with on-line access to the 1975-76 data base located at the application processing site in Iowa City. The CRT was used to access the most recent 1975-76 record of each individual.

Data were extracted manually and were limited to the following items:

- Adjusted gross income
- Signature date
- Whether or not the applicant filed a supplemental form in 1975-76
- Whether the applicant changed his/her dependency status

Data from both years were keypunched and entered into a computer file. The Statistical Package for the Social Sciences (SPSS) computer software was utilized to analyze the data. Analyses were limited to descriptive statistics including measures of central tendency and distribution patterns. Cross-tabulations were performed to answer the analysis questions.

2.3: CRT LOOK-UP STUDY FINDINGS

Findings for the CRT Look-up Study are presented in two sections - the characteristics of the sample, and the findings related to the analysis questions. The subsections which follow present the study results in these two major areas.

2.3.1: Characteristics of the Sample

The sample of 1,547 records had the following characteristics:

- The mean adjusted gross income of applicants was \$5,036 with a range of \$0 to \$38,695 and a standard deviation of \$4,044.
- 1,525 individuals' records included the date of their 1974-75 application. Dates were omitted on 22 applications.
- Of the 1,547 total, 1,127 individuals filed 1975-76 applications. Of these, 1,106 records included the date of the 1975-76 application. 21 records lacked dates.
- 110 or 7 percent of the 1975-76 applicants filed a supplemental report for income estimation.
- 51 or 3 percent of the 1975-76 applicants changed their status from dependent to independent student.
- Final analysis was conducted utilizing 1975-76 applicants who had estimated their 1974 income previously but did not do so on the 1975-76 report and who did not change dependency status. Thus, final comparisons were made on a total of 962 observations.

The findings of this study indicate that 27.2 percent of the individuals estimating their income in 1974-75 who were eligible, on the basis of their supplemental report, did not refile in 1975-76 for a grant. This figure may be attributed to several factors: graduation of students attending institutions with two-year programs; the dropout rate from post-secondary schools; and the possibility that some students may have failed to re-apply if their 1974 income was significantly higher than expected. This might occur because the student does not expect that he/she will qualify.

The 51 individuals who changed their dependency status from dependent to independent between 1974-75 and 1975-76 represent only 3 percent of the total sample. This can be expected, since one of the criteria for filing a supplemental application is the death of a parent. If an applicant was orphaned in 1974, he/she would become an independent in 1975-76.

Supplementals are filed only in the event of an unusual and unanticipated difficulty in the current year. Therefore, a few applicants were expected to file a supplemental report in the second year under investigation. In fact, 110 individuals (or 7 percent of the sample) estimated again in 1975-76. This figure was probably due in part to the small chance occurrence that an individual might have suffered a catastrophe leading to a sudden change in financial status for two consecutive years, or it may have been due to a recurrence of unemployment. In many industries, particularly the automotive industry, furloughs and temporary layoffs are typical and recur year after year. The base years on which the BEOG eligibility was calculated for this study were also years in which unemployment was a serious factor in the American economy. Therefore, even persons who were not employed in industries that typically have layoffs may have suffered unemployment during both years. These explanations probably account for most of the recurrence of estimating, with a small portion of the sample using the mechanism for intentional misreporting, or refiling as a supplemental due to misunderstanding of the filing instructions. However, the exact explanations for the recurrence of supplemental applications for these individuals is not known since the study methodology did not permit follow-up of these individuals to obtain such explanations.

2.3.2: Findings Related to the Analysis Questions

This section presents the findings of the three cross-tabulations conducted to answer the three analysis questions listed in section 2.1. The main objective of the CRT Look-up study was to obtain information related to the size and extent of differences between an individual's estimate of 1974 income for the 1974-75 BEOG application and the actual 1974 income of the individual as reported on the 1975-76 application. The relationships between this difference and 1) the date of filing in 1974-75, 2) the date

of filing in 1975-76, and 3) the income level the applicant over estimated on the 1974-75 application were analyzed in order to determine applicant's ability to estimate.

The difference in income reported on the 1974-75 (estimated) SER from that reported on the 1975-76 (actual) SER was the dependent variable. This variable was broken down into nine (9) categories of discrepancy as listed below (with Difference = Adjusted Gross Income for 1974 reported on 1975-76 application minus the Adjusted Gross Income for 1974 as estimated on the 1974-75 application):

- 1 = Difference $\leq$ -2001
- 2 = Difference $\leq$ -1001, $\geq$ -2000
- 3 = Difference $\leq$ - 501, $\geq$ -1000
- 4 = Difference $\leq$ - 201, $\geq$ - 500
- 5 = Difference $\geq$ - 200, $\leq$ + 200
- 6 = Difference $\geq$ + 201, $\leq$ + 500
- 7 = Difference $\geq$ + 501, $\leq$ +1000
- 8 = Difference $\geq$ +1001, $\leq$ +2000
- 9 = Difference $\geq$ +2001

Thus, categories 1 through 4 represent overestimates; categories 6 through 9 represent underestimates; and category 5 represents insignificant differences.

QUESTION #1: WHAT WAS THE OVERALL ACCURACY OF ESTIMATION?

Of the total of 962 individuals included in the cross-tabulations, 201 (20.9%) overestimated their 1974 income by \$201 or greater; 238 (24.7%) were within \$200 of the actual amount; and the remainder, 523 (54.3%), underestimated. Exhibit 2.1 shows the percentages for each category of discrepancy. Exhibit 2.2 presents this distribution graphically. The largest numbers of individuals occur in the insignificant discrepancy and the over \$2,000 discrepancy categories. The CRT Look-Up Study alone does not provide sufficient evidence as to causes of misreporting or the nature of

EXHIBIT 2.1: DISTRIBUTION OF MISREPORTING FOUND IN CRT-LOOK-UP STUDY

AGI DISCREPANCY BETWEEN 1974-1975 SUPPLEMENTAL AND 1975-1976 APPLICA- TION	NUMBER	PERCENT
<u>Overestimating</u> (subtotal)	(201)	(20.9)
1) -2,000 and less	83	8.6
2) -1,001 - -2,000	41	4.3
3) -501 - -1,000	39	4.0
4) -201 - -500	38	4.0
<u>Accurate</u>		
5) -200 - +200	238	24.7
<u>Underestimating</u> (subtotal)	(523)	(54.3)
6) +200 - +500	53	5.5
7) +501 - +1,000	88	9.1
8) +1,001 - +2,000	127	13.1
9) +2,001 and up	255	26.5
TOTAL	962	100%

misreporting, but, combined with the other component studies, may provide some insight into the characteristics of the applicants or problems with the application that probably affect estimating accuracy.

Some explanations for misreporting follow. The applicants who estimate their income for Basic Grants are a completely different type of population from those who estimate their income for most other types of financial aid. The Parents' Confidential Statement utilized by the Educational Testing Service, for example, instructs all applicants, regardless of their circumstances, to estimate their present year's income. For Basic Grants,

EXHIBIT 2.2: FREQUENCY OF AGI DISCREPANCIES - CRT LOOK-UP STUDY



^{1/} Accurate defined as discrepancy smaller than ± 200 .

only those individuals who have suffered a severe trauma affecting their current year's adjusted gross income may estimate. Of course, there is the possibility for intentional misreporting, but it is likely that inaccurate estimation also results from other factors. For one thing, the effect of the particular circumstance under which an applicant files a supplemental would most likely have a negative emotional impact on the applicant, possibly causing him/her to be pessimistic about his/her ability to obtain income for the remainder of the year. Second, should the need for the applicant to be assured that he/she will be supported adequately for the year may result in a conservative estimate of income so that the BEOG award will be increased. A third, and probably, most prevalent explanation may be that it is difficult to predict the length of time the change in circumstances will affect the individual's income. Fourth, there are many unknowns associated with the future. For example, after being unemployed, the individual may obtain a substantially better job and thus raise the family's income unpredictably.

QUESTION #2: WILL INDIVIDUALS WHO FILE A SUPPLEMENTAL FORM LATER IN THE GRANT APPLICATION YEAR HAVE A MORE ACCURATE ESTIMATE OF THEIR INCOME THAN THOSE WHO FILE EARLIER?

To answer this question, a cross-tabulation between the difference measure and the date of filing the supplemental application was performed. The hypothesis on which this analysis was based was that the later the filing, the more accurate the estimate, since 1) the applicant would have a better idea of his/her past income up to the date of submitting the supplemental application, and 2) the loss of income would have less of a relative effect on the year's income. Data obtained for analysis reflected the availability of information related to the latest information filed by the applicant during the application year. Since this report was the one on which the person's Grant was based, the latest date of filing was the most appropriate. It was expected that most supplemental reports would be filed later in the application year since a

supplemental report is often submitted as a correction to the individual's regular application. Dates of filing were divided into five categories as defined below:

- 1 = Prior to May 1, 1974
- 2 = May, 1974 to July 31, 1974
- 3 = August 1, 1974 to October 31, 1974
- 4 = November 1, 1974 to December 31, 1974
- 5 = After December 31, 1974

Exhibit 2.3 presents the results of this cross-tabulation and Exhibit 2.4 summarizes these results. A Chi-Square test of the significance of the relationships within the table was not performed because, with high numbers of observations, this test will almost always be biased in favor of significance.

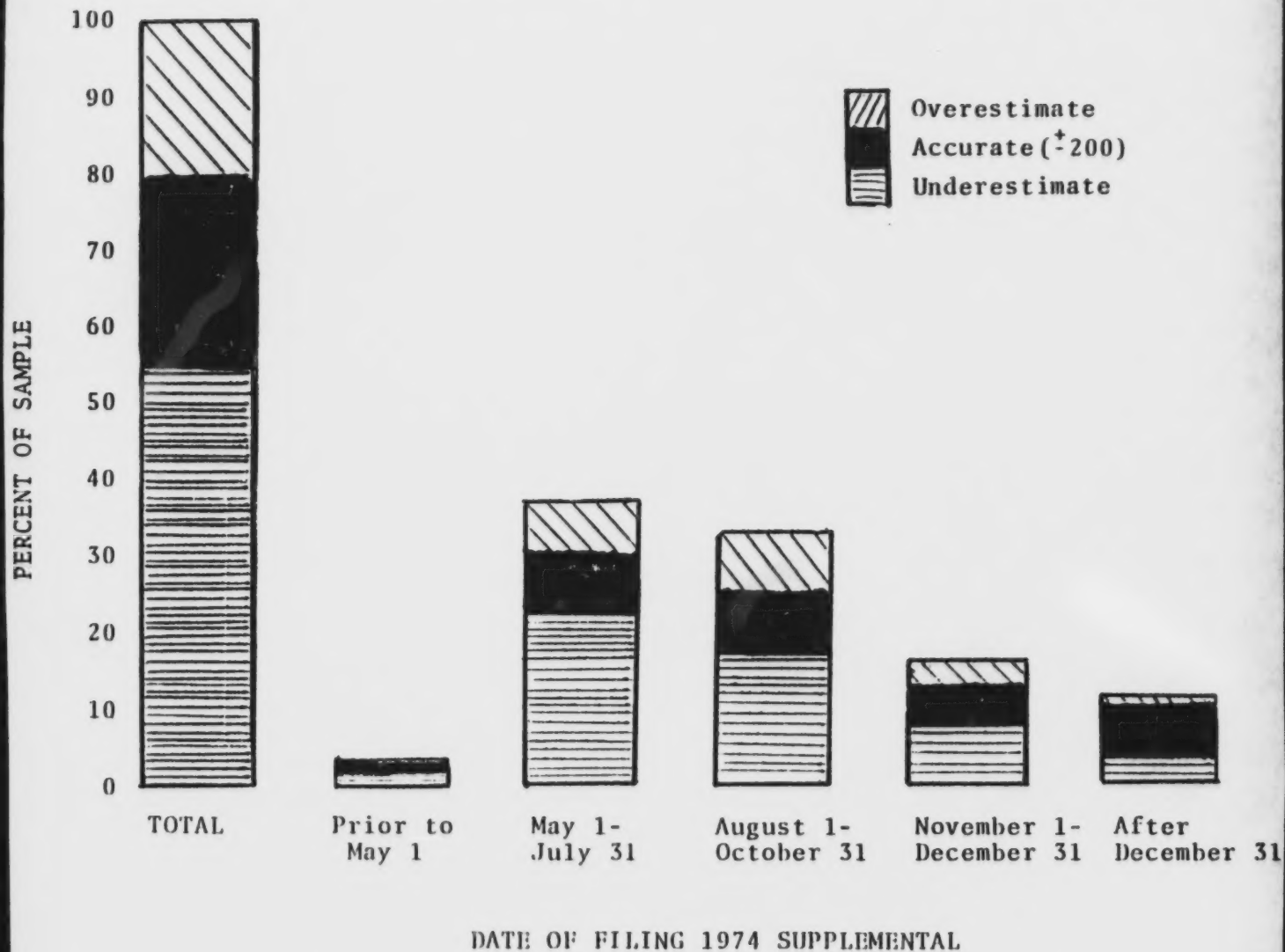
As may be seen in the table, for each data of filing, except those who filed a supplemental application after December 31, the largest groups of applicants either underestimated their income by more than \$1,000 or estimated accurately to within \$200. For those filing after December 31, the percentage of individuals accurately estimating rose substantially. This was expected since these students should have known the total amount earned in 1974.

The table also shows that only 1.2 percent of all persons estimating their income filed a supplemental application before April 30. This was expected, as explained above, since not enough time had elapsed for many of the criterion events for estimating to have occurred, especially unemployment (10 weeks of unemployment is a prerequisite to filing a supplemental). It might also be expected that some time would have elapsed between the time of the individuals' change in circumstances and their date of submission of the supplemental report. This time lag, in turn, may explain the large frequency of estimators in the periods May 1 to July 31 and August 1 through October 31. A smaller frequency was experienced from November 1 to December 31, and after December 31.

EXHIBIT 2.3: DIFFERENCES IN 1974 INCOME BY DATE OF FILING 1974-75 SUPPLEMENTAL APPLICATION

DATE OF FILING ESTIMATE FOR 1974-75		Prior to May 1, 1974	May 1, 1974 to July 31, 1974	August 1, 1974 to October 31, 1974	November 1, 1974 to December 31, 1974	After December 31, 1974	TOTAL	
DIFFERENCE BETWEEN AGI 1974-75 and 1973-74								
1) -2001 and up	1	30	32	13	6	82	COUNT	
	9.1	8.4	10.3	8.6	5.3		COL PCT	
	0.1	3.2	3.4	1.4	0.6	5.7	TOT PCT	
2) -1001 - -2000	1	12	17	3	4	39		
	9.1	3.4	5.4	3.3	3.5			
	0.1	1.3	1.8	0.5	0.4	4.1		
3) -501 - -1000	0	15	15	4	5	39		
	0	4.2	4.8	2.6	4.4			
	0	1.6	1.6	0.4	0.5	4.1		
4) -201 - -500	0	14	14	4	5	37		
	0	3.9	4.5	2.6	4.4			
	0	1.5	1.5	0.4	0.5	3.9		
5) -200 - +200	5	56	63	44	56	234		
	45.5	15.6	20.2	29.1	58.4			
	0.5	5.9	6.7	4.7	6.7	24.8		
6) +200 - +500	0	22	20	5	6	53		
	0	6.1	6.4	3.3	5.3			
	0	2.3	2.1	0.5	0.6	5.6		
7) +501 - +1000	0	33	29	18	7	87		
	0	9.2	9.3	11.9	6.2			
	0	3.4	3.1	1.9	0.7	9.2		
8) +1001 - +2000	2	44	40	30	6	122		
	18.2	12.3	12.8	19.9	5.3			
	0.2	4.7	4.2	3.2	0.6	12.9		
9) +2001 and up	2	132	82	28	8	252		
	18.4	36.2	26.3	18.5	7.1			
	0.2	14.0	8.7	3.0	0.8	26.7		
TOTAL	11	358	312	151	113	945		
	1.2	37.9	33.0	16.0	12.0	100.0		

EXHIBIT 2.4: FREQUENCY OF AGI DISCREPANCIES BY DATE OF FILING 1974 SUPPLEMENTAL APPLICATION



Although fewer applicants filed a supplemental in this period, the reduction in the number of supplementals was greater than that expected, based on the frequency of the preceding and succeeding date categories.

The major discrepancies revealed through this analysis fell in the following areas.

- 64.5% of individuals filing between May 1 and July 1, 1974 underestimated their income.
.. 36.9% of individuals filing in this time period underestimated by more than \$2,000.
- The frequency of accurate filers increased later in the year, especially after December 31. The smallest percentage of accurate filers were in the May 1 to July 31 time period.
- The frequency of underestimators decreased later in the year by roughly the same percentage as the increase in accuracy.
- The percentage of filers in each time period overestimating their income varied by less than 8 percentage points.

In summary, it may be seen that the filing date of an applicant's supplemental SER is related to the difference between estimated and reported income for 1974.

QUESTION #3: HOW WILL THE DATES OF FILING 1975-76 APPLICATIONS BE DISTRIBUTED FOR APPLICANTS WHO HAD FILED A SUPPLEMENTAL FORM IN THE PREVIOUS YEAR?

A cross-tabulation of the differences between 1974-75 income estimates and 1975-76 income reports -- the date of filing the 1975-76 application -- was performed. This analysis was conducted to explore the relationship of the accuracy of estimates to the date of filing the 1975-76 application. Data reflect the availability of information related to the last SER filed by the applicant during the 1975-76 Grant year. Since this SER was the one on which the individual's grant is based, the latest date of filing was the most appropriate. Dates of filing were divided into three categories as defined below:

- 1 = Prior to March 15, 1975
- 2 = March 16, 1975 to April 30, 1975
- 3 = After April 30, 1975

Exhibit 2.5 presents the cross-tabulation performed. Exhibit 2.6 illustrates this distribution. As may be seen in the table, the frequency of accurate estimating was fairly consistent among the 1975-76 filing dates, with 24.6 percent of all individuals filing before April 30 having no significant difference in income between the estimated and reported amounts. Slightly more, 26.6 percent, of those filing after April 30 had no significant differences in their applications. Further, it would appear that there was little variability in discrepancies across the 1975-76 filing dates.

The most notable result of this analysis is illustrated in the fact that 74 percent of the 1974-75 estimators who filed again in 1975-76 did so prior to April 30 of the second Grant year. In the regular filing population, only 17 percent of the 1974-75 eligibles filed prior to April 15. This may be explained partially by the fact that the regular applicant population includes first-time as well as former applicants, while all of the applicants in this study were by definition, former applicants.

Since the proportion of discrepancies were consistent across application dates, it does not appear that people who estimated in 1974 filed early and reported the same amount on their 1975-76 application as was estimated on the 1974 supplemental application. Rather, it would appear that students who refiled applications reported, in the next Grant year, their actual income for the estimated year. It was, of course, also possible that the amount reported on the 1975-76 application is discrepant. See Chapter 3 for a discussion of this issue.

To summarize, it would appear that there was no significant difference in accuracy rates which was related to the date of filing in the year after the estimation procedure was used. It was found, however, that applicants who estimated their income and qualified for an award in one year tended to file their next year's applications quite early in the year.

EXHIBIT 2.5: FREQUENCIES OF DISCREPANCIES IN 1974 INCOME BY DATE OF FILING 1975-76 APPLICATION

AGI DISCREPANCIES	DATE OF FILING 1975-76 APPLICATION				TOTAL
	Prior to March 15, 1975	March 16, 1974 to April 30, 1974	After April 30, 1974		
OVERESTIMATION	81	55	58		194
	18	22.5	23.8		
	8.6	5.9	6.2		20.4
ACCURATE (+200)	111	60	65		236
	24.6	24.6	26.6		
	11.8	6.4	6.9		25.2
UNDERESTIMATION	258	129	121		508
	57.3	52.9	49.6		
	27.5	13.8	12.9		54.2
TOTAL	450	244	244		938
	48	26	26		100

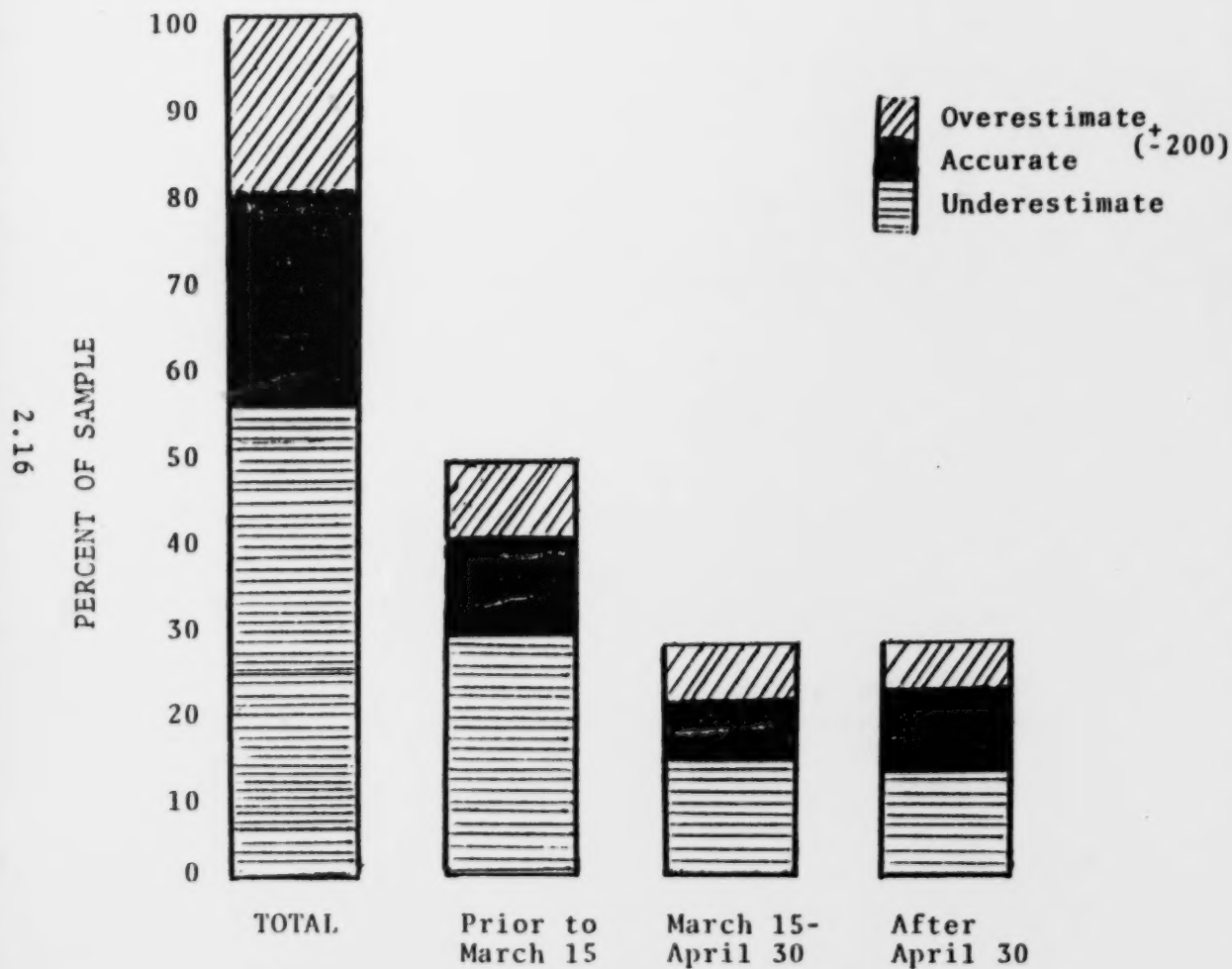
Count

Column Percent

Total Percent

2.15

EXHIBIT 2.6: FREQUENCY OF AGI DISCREPANCIES BY DATE OF FILING 1975-76 BEOG APPLICATION



QUESTION #4: HOW IS THE INCOME LEVEL OF THE APPLICATION RELATED TO THE ACCURACY OF ESTIMATION?

A cross-tabulation was performed between the accuracy of the income estimate and the amount of the income estimate. It was hypothesized that discrepancies would increase as the AGI increased, due to the difficulty of estimating. A discrepancy of \$1,000 for a family with an AGI of \$10,000 is proportionately smaller than if the family with a \$4,000 AGI had the same \$1,000 discrepancy. Discrepancies were defined, as in other analyses, as the AGI reported for 1974 on the 1975-76 application minus the AGI estimated for 1974 on the 1974-75 Basic Grant application. The adjusted gross income was defined as the AGI estimated on the 1974-75 application. Exhibit 2.7 presents the findings of the cross-tabulation. Exhibit 2.8 illustrates the frequency of discrepancies. The AGI variable was divided, for analysis purposes, into the following categories:

- 1 = 0 - \$3,999
- 2 = \$4,000 - \$6,999
- 3 = \$7,000 - \$9,999
- 4 = \$10,000 - \$12,499
- 5 = \$12,500 and up

While it must be remembered that the 1974 estimate was not a completely valid measure of income, as illustrated by the occurrence of discrepancies, it did provide a perspective on the incidence of poor estimating. For example, over 62 percent of individuals estimating accurately had AGIs, as estimated, of under \$4,000. As the income increased, however, the accuracy rates generally diminished substantially. Significant findings are summarized as follows:

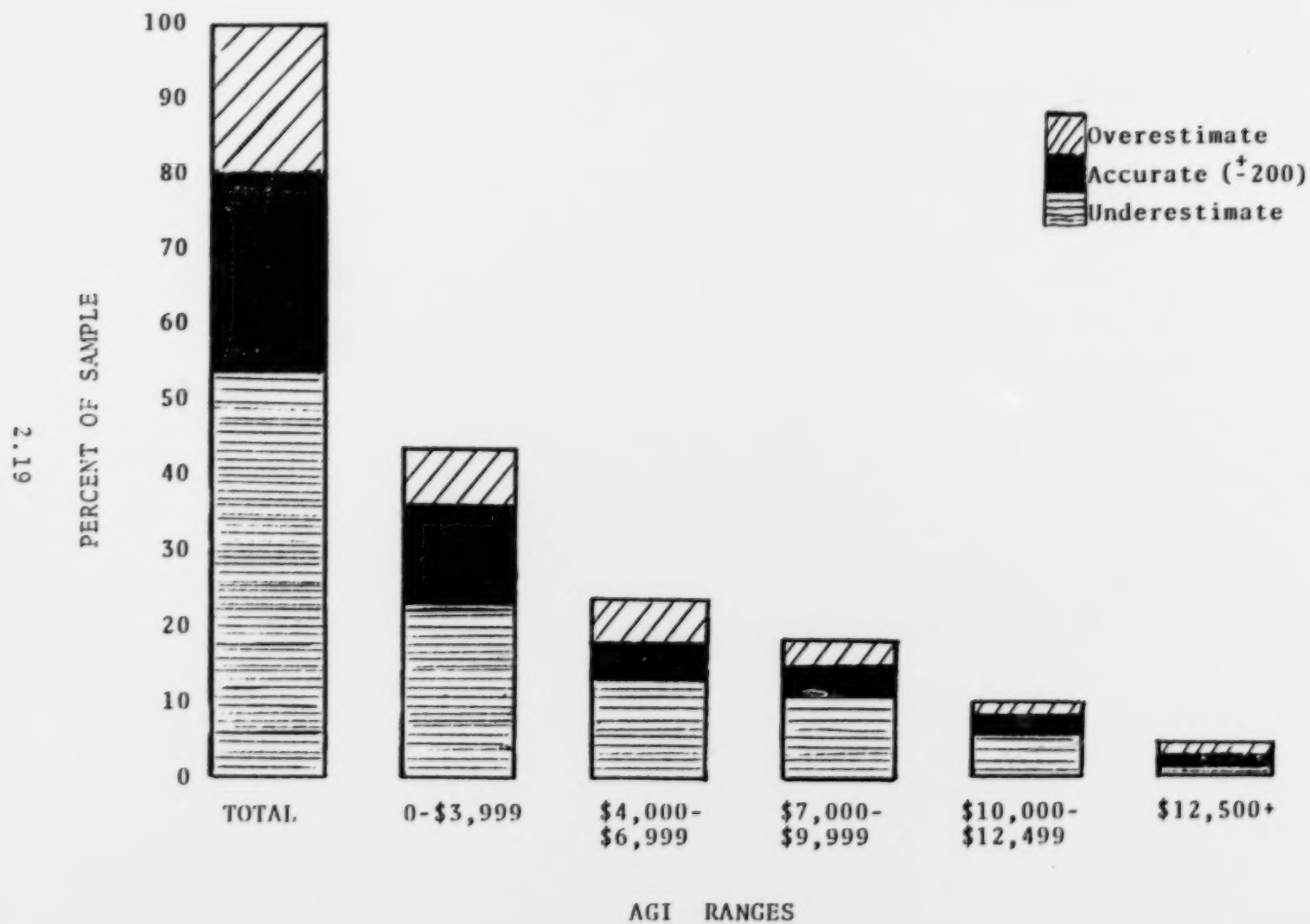
- 43.8 percent reported their 1974 income to be less than \$4,000. 86 percent of the sample estimated their income to be less than \$10,000 in 1974. Less than 1 percent of the sample had estimated incomes of over \$15,000.

EXHIBIT 2.7: AGI DISCREPANCIES BY ADJUSTED GROSS INCOME
LEVEL

AGI ESTIMATED IN 1974-'5 AGI DISCREPANCY	AGI DISCREPANCY								COUNT	COL PCT TOT PCT
	0 - \$3,999	\$4,000 - \$6,999	\$7,000 - \$9,999	\$10,000 - \$12,499	\$12,500 - \$14,999	\$15,000 - \$17,499	\$17,500 - up	TOTAL		
1) -2001 or less	15 3.6 1.6	28 12.1 2.9	22 12.6 2.3	12 13.0 1.2	5 15.2 0.3	1 14.3 0.1	0 0 0	83		
2) -1001 - -2000	12 2.9 1.2	15 6.3 1.6	4 2.3 0.4	6 6.3 0.6	3 9.0 0.3	1 14.3 0.1	0 0 0	41		4.3
3) -501 - -1000	9 2.1 0.9	11 4.7 1.1	10 5.7 1.0	4 4.3 0.4	5 15.2 0.3	0 0 0	0 0 0	39		4.1
4) -201 - -500	17 4 1.8	13 5.6 1.4	3 1.7 0.3	4 4.3 0.4	1 3 0.1	0 0 0	0 0 0	38		4.0
5) -200 - +200	148 35.2 15.4	41 17.7 4.3	30 17.2 3.1	13 14.1 1.4	3 9 0.3	1 14.3 0.1	2 66.7 24.7	238		
6) -200 - +500	23 5.5 2.4	16 6.3 1.7	9 5.3 0.9	4 4.3 0.4	1 3 0.1	0 0 0	0 0 0	53		5.3
7) -501 - +1000	35 8.3 3.6	20 9.6 2.1	20 11.5 2.1	9 9.8 0.9	3 9 0.3	0 0 0	1 33.3 0.1	38		9.1
8) +1001 - +2000	46 10.9 4.5	30 12.9 3.1	28 16.1 2.9	17 19.3 1.8	5 15.2 0.3	1 14.3 0.1	0 0 0	127		13.2
9) +2001 or more	115 27.6 12.1	58 25 6	48 27.6 5.0	23 25 2.4	7 21 0.7	3 42.9 0.3	0 0 0	255		26.3
TOTAL	421 45.3	332 24.2	174 19.1	92 9.6	33 3.4	7 7	3 3	962 100		

BEST COPY AVAILABLE

EXHIBIT 2.8: FREQUENCY OF AGI DISCREPANCIES BY ADJUSTED GROSS INCOME LEVEL



- Accurate estimators accounted for 35 percent of applicants in the lowest income category although only 24.7 percent of all applicants had insignificant discrepancies. In addition, persons in this category tended to underestimate their income much less frequently than those in other income categories.
- In the \$4,000-\$6,999 income category, a different pattern arises. 28.9 percent of individuals in this category over-estimated their income with only 17.7 percent estimating accurately. Over 50 percent underestimated.
- Both over and under-estimating occurred more frequently as the AGI goes over \$4,000. For larger incomes, the larger differences (over \$1,000) occurred more frequently, generally in the direction of underestimating. As AGI increases, these differences appeared to reflect a smaller percentage of the individual's income.
- In the over \$12,500 category, the number of over-estimators increased and the number of underestimators declined somewhat.

A discussion of these findings may be found in section 2.4. In general, estimators with AGIs of less than \$4,000 had a high accuracy rate and a high rate of underestimation. As AGI increased the accuracy rate declined sharply. Thus, accuracy was less than the expected frequency at all but the lowest income levels. This may be explained to some extent by the fact that the lowest range included the zero income level and may have accounted for negative AGIs as well. Although this category includes 44 percent of all applicants, it includes 62 percent of all accurate estimators. Thirty-five percent of all persons in the under \$4,000 income category are accurate estimators as opposed to 24.7 percent for all applicants. This income category did account for an expected number of underestimators, with 38.5 percent of its members underestimating by \$1,000 or more. It should be noted that, in this group, an error of the magnitude of \$1,000 was a greater percent of total income as estimated than for other income groups. It follows, then, that less than the expected number in this category overestimated their income. In fact 3.6 percent overestimated with 8.6 percent (as reflected in the distribution for the total sample) expected to do so.

In the next category of AGI, \$4,000-\$6,999, far fewer individuals were accurate. Here, overestimation accounted for 28.9 percent of persons in this category with 17.7 percent estimating accurately. It might be noted that over 50 percent again underestimated. The frequency of estimating accuracy appeared to be affected by the level of income, although there is a high amount of inaccuracy at all levels.

2.4: CRT LOOK-UP SUMMARY AND CONCLUSION

The CRT Look-up study was designed to compare the estimates of 1974 income by persons who were permitted to file a supplemental in 1974-75 with the regular applications filed by these individuals in 1975-76. Since the 1975-76 regular applications were based on a report by the applicants of their actual 1974 income, this comparison was able to provide the following information about the accuracy of estimating:

- Accuracy was related to the estimated income level of the applicant. Those with lower incomes were more accurate; those with higher incomes had a higher incidence of overestimating.
- Accuracy was related to the date of filing the estimated application. The later in the year the supplemental was filed the more accurate the estimate.
- Accuracy was not related to the date in which the second year's application was filed.
- As the rate of accurate estimates increased, the size as well as the frequency of discrepancies became smaller.

Overall, accurate estimates occurred in only about one quarter of the cases. Most estimates were more than \$2,000 different from what the applicants later reported that they earned in 1974. Because the criterion of accuracy in the CRT Look-up Study was the 1975-76 application of each sample member (where actual 1974 income may have been reported erroneously), the amount of accurate reporting probably was overstated. The Individual Case Follow-up component of this study addressed this issue and found 1975-76 applications to be significantly discrepant from IRS data provided by its sample members.

However, other study components confirmed the finding of relationships to income levels as estimated and to the latest signature date of applicant's records.

3

INDIVIDUAL CASE FOLLOW-UP

The Estimated Income Individual Case Follow-up Study was designed to determine the accuracy of income estimates provided by BEOG applicants. In this component of the Estimated Income investigation, information verifying their estimated income was obtained from a sample of eligible applicants. Verification utilized a notarized copy of the individual's or his/her family's income tax return (IRS Form 1040 or 1040A). The Estimated Income Individual Case Follow-up procedures generally paralleled those of the Pre-Established Criteria study with one exception - no attempt was made in this effort to have the applicants change their SERS or to adjust the students' award in any way. This was done for two important reasons. First, without knowing if a poor estimate was intentionally made, it must be assumed that the estimate was determined by the applicant in good faith. Second, estimates are allowed only in unusual circumstances. These circumstances were likely to carry considerable emotional as well as financial problems for the applicant. Since the Estimated Income Study is primarily aimed at studying the accuracy of estimates, it was felt that data collection was the most important aspect of the effort. Thus, in order to obtain the best statistical data for the purposes of the study and to encourage response, no further attempt was planned or implemented for altering the estimates. The IRS 1040 or 1040A income tax return was selected as the criterion of accuracy in this study.

The Individual Case Follow-up Component complements the CRT Look-up by examining a different criterion document (IRS 1040 or 1040A as opposed to the 1975-76 BEOG application). At the same time, the Individual Case Follow-up methodology allowed a more complete picture of the estimating population for two reasons. First, the reasons for filing the supplemental applications could be analyzed in this study. Second, the case follow-up procedure allowed direct contact with members of the sample, increasing substantially the potential for background information on this population.

3.1: STUDY OBJECTIVES AND ANALYSIS QUESTIONS

The Individual Case Follow-up component of the Estimated Income Study provided information concerning the accuracy of estimations and characteristics of applicants who estimated their incomes. This information is related to each of the main study objectives listed in Chapter 1. Specifically, analysis questions for which Individual Case Study data were collected are:

- 1) What is the overall accuracy of estimation?
- 2) Is accuracy of estimation related in any way to the reasons for which the supplemental was filed?
- 3) How is accuracy of estimation related to the applicant's adjusted gross income as reported to Basic Grants?
- 4) Is the discrepancy between 1974-75 income estimated to BEOG and income reported on 1974 Federal Tax Return related to a discrepancy between actual 1975-76 income reported to BEOG and reported on 1974 Federal Tax Return?
- 5) Are the IRS forms submitted to the validation contractor true copies of those submitted to IRS for tax purposes?

Question #1 investigates the frequency and extent of misreporting. Question #2 expands the investigation of the characteristics of inaccurate estimators beyond income and date of filing characteristics looked into by the CRT Look-up Study. At the same

3.2

time, Question #3 provides further data on income characteristics. Question #4 will allow an analysis of the usefulness of the 1975-76 application as the criterion of accuracy in the CRT Look-up Study as well as providing further data on misreporting.

Question #5 will not be discussed in this chapter because it required comparison with IRS Sample data presented in Chapter 4, and will be discussed there.

The findings and conclusions drawn from analysis specifically related to the Individual Case Follow-up Study will be presented in this chapter. The more comprehensive study objectives require data from all three studies in order to be answered. These will be discussed in Chapter 5. The remainder of the present chapter describes the sampling procedures, methodology, findings and conclusions of the Estimated Income Individual Case Follow-up Component.

3.2 INDIVIDUAL CASE FOLLOW-UP METHODOLOGY

The Individual Case Follow-up component of the Estimated Income Study utilized a stratified random sample of estimators from the BEOG 1974-75 master file who also qualified for sampling in the IRS Study. These individuals were contacted by mail and asked to supply verifying information for their BEOG-reported income data.

3.2.1: Individual Case Follow-up Sampling Procedures

A sample of 408 individuals who estimated their 1974 income on the 1974-75 Basic Grant application were selected for follow-up contact. The sample was drawn from a population defined as those individuals who:

- filed a supplemental application and estimated the 1974 income;
- were eligible for a Grant; and
- if dependent, whose parents' social security numbers were available.

A stratified random sample was utilized based on BEOG-reported income. A skip interval with random start procedure was implemented for selection. The strata, population and sample sizes are defined below:

Strata	Population Count	Approximate Sample Size
\$ 0 - \$ 3,999 income	955	104
\$4,000 - \$ 9,999 income	1,551	179
over \$10,000 income	1,023	125
TOTAL	3,529	408

This sampling frame was based on a requirement by the Office of Education for a sample size of 400. Of the 408, 58 had not submitted an acceptable response by the date of analysis (which was June 6, 1976). An additional 14 individuals could not be located by the validation staff. For each category of application (as defined by the reasons for which they submitted supplemental forms) missing cases accounted for 1 in 6 sample members, except in the case of those reporting "separated or divorced" where this rate was approximately 1 in 3 (31%).

3.2.2: Individual Case Follow-up Methodology

The Individual Case Follow-up involved contacting 1974-75 supplemental applicants to request that they submit documentation of their actual 1974 income as reported to IRS. This information was compared to the Adjusted Gross Income entry on the student's SER in an effort to determine the degree of accuracy generally obtained in income estimates. No action was taken vis a vis changing an applicant's eligibility index or adjusting his/her grant. Accordingly, respondents were informed that the objective of the study was only to determine the level of accuracy and that no changes would be made to the Basic Grant award based on the applicant's response.

An initial contact letter was mailed out to students on December 17, 1975. If no response was received after one month, a follow-up letter was sent. After another month a second follow-up letter was sent to those individuals who had not responded either by telephone or letter. If no response was obtained within another month, these cases were considered closed due to "non-response." Cases where initial or subsequent contact letters were returned as undeliverable by the U.S. Postal Service were closed and categorized as "unable to contact student/parents."

As responses were received a case report form was completed for each BEOG applicant, and the response was checked for acceptability. Criteria for the acceptability of verification information were presence of one of the following:

- 1) Notarized IRS-1040 and W-2 forms;
- 2) Letter indicating AGI if income was too low or otherwise did not require a 1040 or 1040A to be sent to IRS;
- 3) IRS-1040 or 1040A and W-2 not notarized if the tax data were prepared by an outside agent;
- 4) IRS-1040 or 1040A not notarized if signature date was prior to April 15 and the AGI and taxes differed from the estimate;
- 5) Notarized W-2 forms if no IRS 1040 or 1040A were available.

The respondents were not given the options included in criteria 2 through 5 unless he/she contacted the validation staff to say that it was not possible to obtain the requested information.

After the acceptability of the information was determined, the applicant was sent an acknowledgement letter, and the response was processed. The AGI from the documentation was compared with the BEOG reported data, and the amount of the difference was recorded on the individual's case report form. Next, the 1975-76 master file was checked to determine whether an application was filed in 1975-76. If the 1975-76 file indicated that the

applicant had filed a regular 1975-76 application, the AGI reported on the 1975-76 application was compared with the 1974-75 BEOG application estimate and the AGI reported to IRS. The amount of difference and the date the 1975-76 application was filed were recorded on the individual's case report form.

Data were keypunched from the Individual Case Report Forms for all closed cases. The information required for analysis included the following data elements;

- Adjusted Gross Income reported on BEOG application
- Adjusted Gross Income reported on IRS Form 1040 or 1040A, or other documentation provided
- Reason for filing supplemental
- Discrepancy between BEOG and IRS data
- Transactions made

Analysis involved cross tabulations of the data. Analysis of transactions was performed to determine the usefulness of processing methods and is discussed in Volume I of this report. Applied Management Sciences' generalized table production program (PROTABS) was utilized to develop the tables.

3.3: INDIVIDUAL CASE FOLLOW-UP FINDINGS

Comparative data from both 1974-75 estimated income applications and the 1974 IRS 1040 or 1040A were available for 336 of the 400 sampled. All of these 336 individuals were included in the analysis. The following subsections present the characteristics of the sample and the results of the analyses.

3.3.1: Characteristics of the Sample

The following table shows the initial sample and the number and percent of usable responses received:

	Initial Sample	Usable Responses
Total	408	336 82.3%
Supplemental Reason:		
Death	58	49 84.4%
Loss of employment	180	150 88.3%
Unable to pursue income activities	144	119 82.6%
Separation or divorce	26	18 69.2%

As is evidenced from this exhibit, 82.3 percent of the applicants surveyed provided a usable response. In terms of reason for filing a supplemental, this response rate did not vary, except for applicants who cite divorce or separation as the reason for filing a supplemental. Usable responses were only received from 69 percent of this group; however, divorce and separation only account for 6 percent of the supplemental applications filed. Consequently, the lower response rate for this group does not impact very greatly on the overall findings.

Some of the major characteristics of the respondent sample may be summarized as follows:

- Reason for filing supplemental
 - .. 14.6% cited death
 - .. 44.6% cited loss of employment
 - .. 35.4% cited unable to pursue income activities
 - .. 5.4% cited separation or divorce
- Actual income levels as defined by IRS form or other documentation provided
 - .. 25.3% under \$4,000
 - .. 30.7% between \$4,001 and \$9,999
 - .. 44.0% over \$10,000

Records of 72 individuals in the sample of 408 were not included in analysis for various reasons. Several were cases of non-response. Other cases were closed by the Office of Education. The causes of missing cases are explained as follows:

- During the course of the study, several individuals sent letters of explanation concerning their sons/daughters who qualified for but did not receive a Grant. These individuals did not see any reason for supplying the information requested. Project staff contacted these persons by telephone and explained to them the nature of the study and the need for the information requested. During this contact, most individuals made a verbal commitment to send the information. However, by the end of the project it was discovered that many of these persons had not sent the documentation.
- A group of individuals contacted the Validation Study staff to indicate that they were in the process of sending the documentation but required more than the two weeks OE suggested in the contact letters. At the end of the study any cases which had had such contact but for which no documentation was as yet received were considered, for the purposes of this report only, to have been closed, although actual case processing continued to the end of the contract.
- Several other cases were closed by decisions made by the Office of Education. These cases involved individuals for whom circumstances were unusual and whose data would not be useful. For example, OE closed one case where the BEOG recipient had passed away soon after filing the supplemental form.
- Some applicants' data were not included, by OE decision, because of difficulty of interpretation. These consisted of cases where the parents of the student were divorced, but had filed a joint tax return in the year of the divorce. In this event the only way of determining the amount of income of that parent who supplied more than 50 percent of the student's support, would have been through a telephone contact of the parent. OE directed that these cases not be followed up because of the sensitive nature of the family situation.

3.3.2: Findings Related to the Analysis Questions

This section presents the findings of cross-tabulations performed to answer several of the analysis questions listed in section 3.1. The major purpose of the Individual Case Follow-up Study was to obtain information related to the size and extent of the difference between an individual's 1974-75 estimated BEOG application and that individual's income tax return as supplied by her/him to Basic Grants. First, the overall accuracy of the sample was examined. Two of the analysis questions treated the nature of misreporting by income level and reason for filing a supplemental form. Another looked at the misreporting of persons whose 1975-76 application was also discrepant. The fifth question is addressed in Chapter 4 in conjunction with IRS Sample results.

The difference in income reported on the 1974-75 supplemental SER from that contained in a copy of the applicant's 1974 Federal Income Tax Return was the criterion of accuracy for this study. The categories of discrepancy vary^{1/} somewhat from those utilized in Chapters 2 and 4 but comparisons with the conclusions of other components can be readily performed. The categories used for this component are:

- 1 = Difference < -3000
- 2 = Difference $\geq$ -3000, $\leq$ -1001
- 3 = Difference $\geq$ -1000, $\leq$ -501
- 4 = Difference $\geq$ -501, $\leq$ -301
- 5 = Difference $\geq$ -300, $\leq$ +300
- 6 = Difference $\geq$ +301, $\leq$ +500
- 7 = Difference $\geq$ +501, $\leq$ 1000
- 8 = Difference $\geq$ +1001, $\leq$ 3000
- 9 = Difference $\geq$ +3000

^{1/} In the CRT look-up Study and IRS Comparison Study, accuracy was defined as no discrepancy or a discrepancy of + \$200. In this study, + \$300 was the range for accuracy. The reason for this inconsistency is that the Individual Validation Analysis plan was specified prior to the CRT and IRS analysis plans. Subsequent to the Individual Validation Study Analysis plan, OE decided to use the more stringent definition for accuracy. Consequently, the findings of the three studies are not totally comparable.

Categories 1 through 4 represent overestimating, category 5 represents either no difference between estimating and actual income or no significant difference; and categories 6 through 9 representing underestimating.

QUESTION #1: WHAT IS THE OVERALL ACCURACY OF ESTIMATION?

The following observations can be made about the accuracy of estimation in the Individual Follow-up Study:

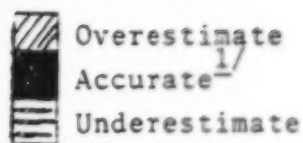
- 61.6 percent of the sample underestimated their income on their BEOG application compared to their IRS return. In the CRT Look-up Study, 52.8 percent underestimated by more than \$300.
.. 25.6 percent were found to have underestimated by more than \$3,000, while only 17.2 percent of the CRT Look-up Study sample showed underestimations of \$3,000 or more.
- 17 percent estimated accurately (within $\pm$ \$300). This may be compared to the 28.5 percent in this discrepancy range in the CRT-Look-up analysis.
- 21.4 percent overestimated their 1974 income. This is slightly higher than for the CRT Look-up.

Another interesting finding related to the rate of accuracy is that while the proportion of persons sampled included 30.6 percent who had estimated incomes over \$10,000; the proportion of the sample with actual AGIs that large according to the documentation supplied was 44 percent, indicating substantial amount of discrepancies. This is discussed further in response to Question #3.

QUESTION #2: IS ACCURACY OF ESTIMATED INCOME RELATED IN ANY WAY TO THE REASONS FOR WHICH THE SUPPLEMENTAL WAS FILED?

In order to investigate this question, a cross-tabulation was performed on the parameters of the discrepancy data as explained above and the four criteria for filing supplemental SERs. See Exhibits 3.2 and 3.3. The overall finding of this analysis is that, in general, the degree of accuracy does not vary greatly across reasons for filing a supplemental. That is, in most cases the variability of the proportion of discrepancies across reasons for filing was small. Specific observations concerning these data may be summarized as follows:

EXHIBIT 3.1: FREQUENCY OF AGI DISCREPANCIES - INDIVIDUAL
CASE FOLLOW-UP STUDY

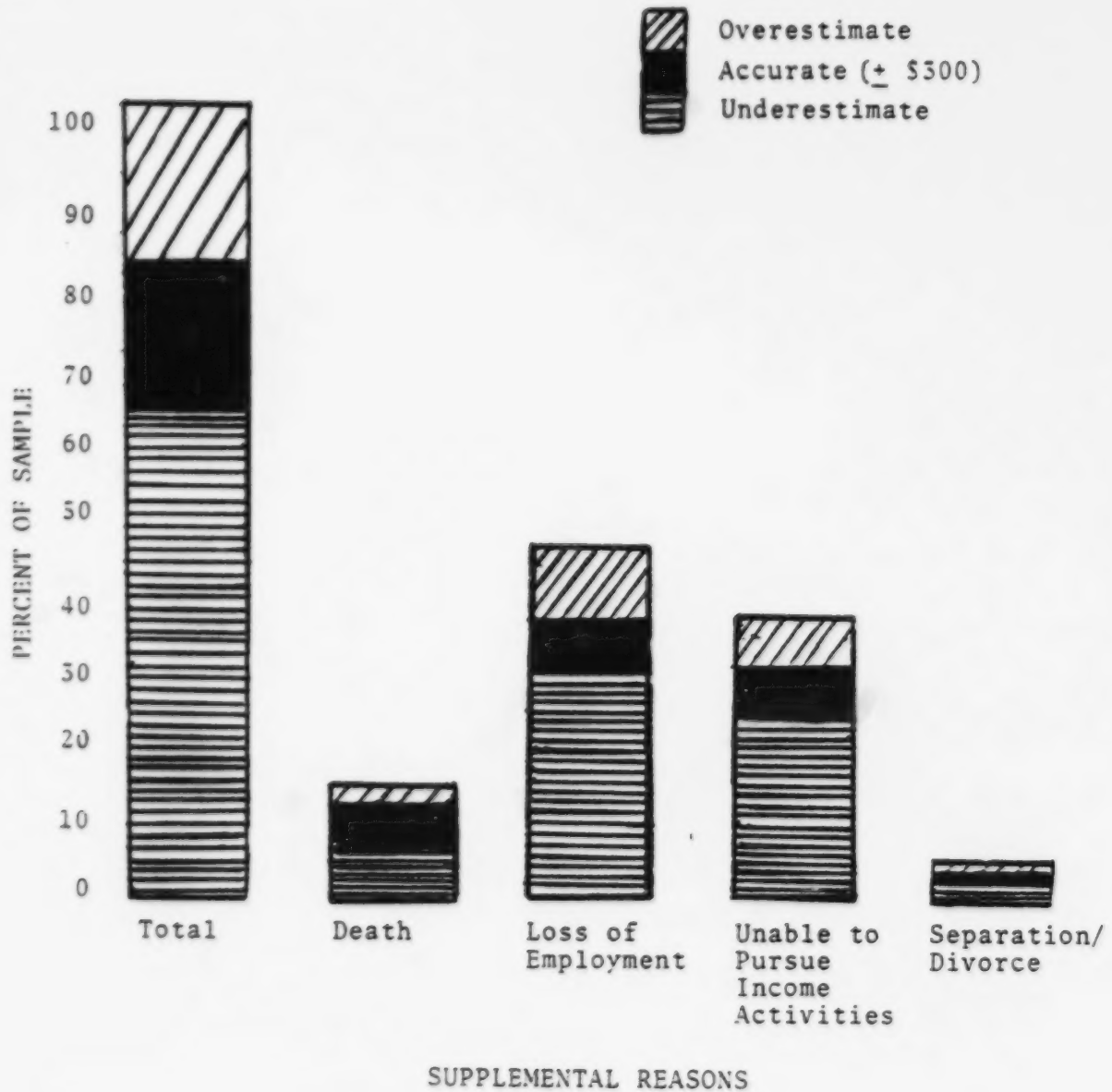


^{1/} Accurate is defined as a discrepancy less than $\pm$ 500.

EXHIBIT 3.2: REASON FOR FILING SUPPLEMENTAL BY SIZE OF
DISCREPANCY IN 1974 INCOME

AGI DISCREPANCY	Reason for Filing Supplemental					COUNT COL PCT TOTAL PCT
	DEATH	LOSS OF EMPLOYMENT	UNABLE TO PURSUE INCOME ACTIVITIES	SEPARATION OR DIVORCE	TOTAL	
1) -3,000 or less	6 12.2 1.3	4 2.7 1.2	11 9.2 3.3	2 11.1 0.6	23 6.8	
2) -3,000 - 1,001	2 4.1 0.6	13 8.6 3.9	8 6.7 2.3	3 16.7 0.9	26 7.7	
3) -1,000 - 501	2 4.1 0.6	12 8. 3.6	3 2.5 0.9	0 0 0	17 5.1	
4) 500 - 301	2 4.1 0.6	2 1.3 0.6	2 1.7 0.6	0 0 0	6 1.8	
5) -300 - +300	12 24.5 3.6	21 14 6.3	20 16.8 6.0	4 22 1.2	57 16.9	
6) 301 - 500	1 2. 0.3	6 4 1.8	9 7.5 2.6	1 5.5 0.3	17 5.1	
7) 501 - 1,000	2 8. 4.1 0.6	13 52. 8.7 3.9	9 36 7.6 2.7	1 4 5.5 0.3	25 7.4	
8) 1,001 - 3,000	10 20.4 3.0	40 26.7 11.9	27 22.7 6.0	2 11.1 0.6	79 23.5	
9) More than 3000	12 14. 24.5 3.6	39 26. 11.6	30 25.2 8.9	5 27.8 1.5	86 25.6	
TOTAL	49 14.6	150 44.6	119 35.4	18 5.4	336	

EXHIBIT 3.3: FREQUENCY OF AGI DISCREPANCIES BY
SUPPLEMENTAL REASON



- Persons citing death as the reason for the supplemental were most accurate (24%). Members of this group also had significant underestimating (44.9%), although this was less than for other groups. A higher percentage in this group also overestimated significantly.
- A similar pattern is evident for cases of divorce/separation. However, this group is not a reliable gauge of trends for several reasons:
 - .. Size - it is the smallest in the sample with only 5.4 percent of the population
 - .. Response Rate - response rate in this group was the poorest and thus, the validity of responses most in question
 - .. Because many divorced parents file joint returns, it was difficult to determine the contribution of the income of the parent providing the major support to the student, thus confounding the information
 - .. It is easiest for an applicant to file a false supplemental for this category than for the others because there need not be any legal agreement, and, further, the year's tax return may not necessarily contain evidence about a separation.
- The categories for unemployment and unable to pursue income activity had similar distributions of discrepancies in estimation, but more heavily concentrated in underestimation.

The larger discrepancies for the latter two categories may be due to the applicant's inability to predict accurately the end of the no-income period and/or salary levels after work is resumed. In cases of death and separation/divorce, a more finite perception of future income may be available. However, differences among the groups in percent of accurate reporting did not vary by more than 10 percent. Because such large portions of each applicant group underestimated income, it does not appear that misreporting is strongly associated with the reason for filing a supplemental.

QUESTION #3: HOW IS THE ACCURACY OF ESTIMATION RELATED TO THE APPLICANT'S IRS-REPORTED AGI?

The criterion for the income and accuracy in this analysis is the AGI contained in a copy of the applicant's or parents' 1974 Federal Income Tax Return. It is assumed that an individual reports his/her true income to IRS. However, it may not be the case that the applicant, in complying with the request for verifying information, provided a true copy of the IRS form 1040 or 1040A to the validation contractor. This latter condition is investigated in Chapter 4.

The actual AGI contained in the requested documents was cross-tabulated with the discrepancy level of the individual's BEOG estimated income. It should be noted here that a similar analysis in the CRT Look-up Study was based on the level of the estimated AGI, not the actual income contained in a copy of the tax form submitted by the applicant. Generalizations to be made from the two analyses will be somewhat affected by this difference.

Exhibit 3.4 summarizes the results of the cross-tabulation; Exhibit 3.5 illustrates these findings. In the category of persons with true low incomes, comparatively more individuals over-estimated their income to BEOG, while the opposite was true for persons with high incomes. Applicants in the middle income category were not significantly different in their degree of accuracy from the total sample. It should be noted that this analysis was somewhat biased against accuracy, because of the parameters used to define accuracy. Here, accuracy was defined as an exact match so that patterns of over- and underestimating could be more clearly distinguished by income level. Analyses performed in response to Question 2, however, considered discrepancies to be greater than an absolute difference of \$300 so that the overall scope of inaccurate estimation could be assessed meaningfully.

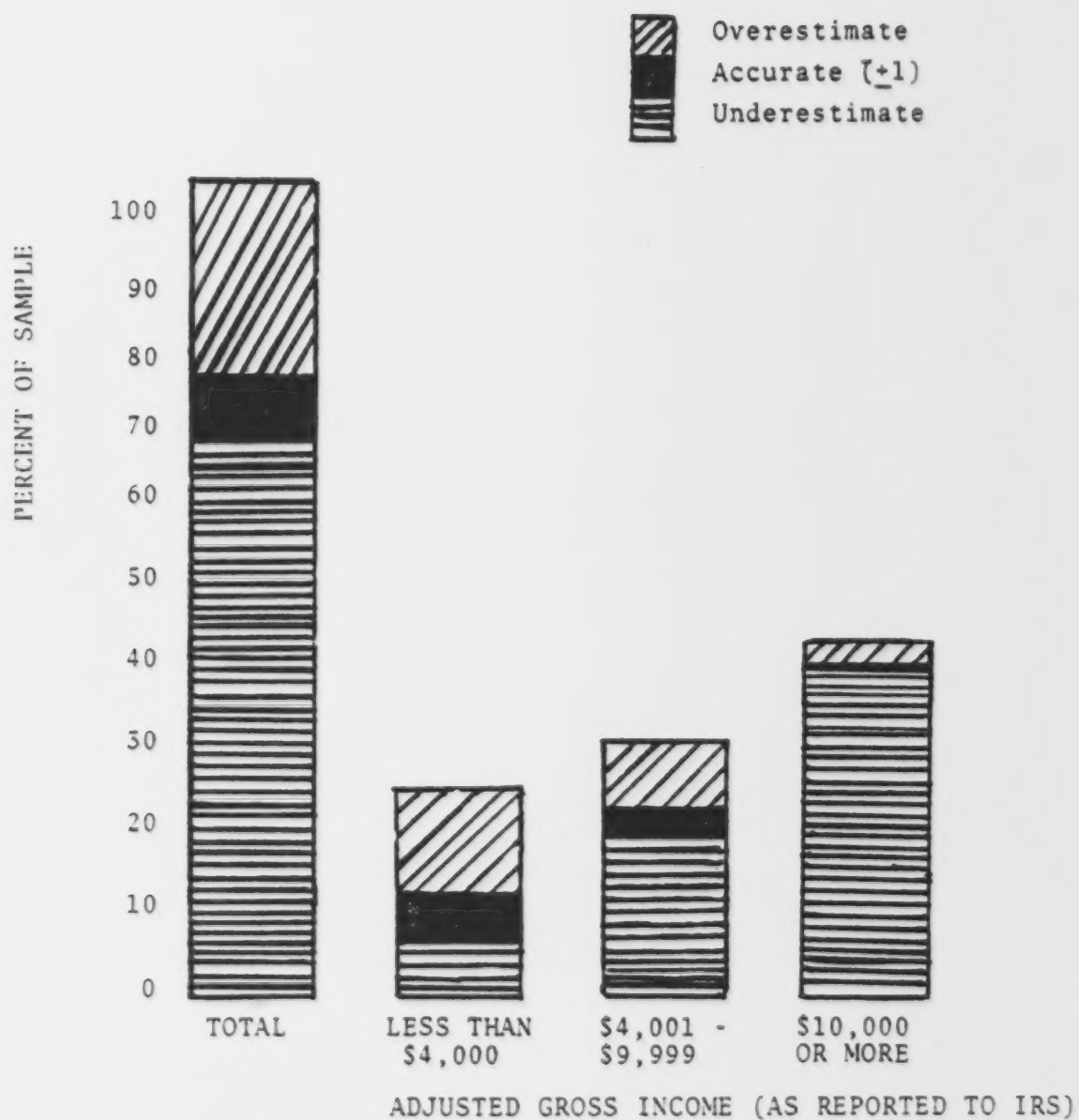
Significant variations in the accuracy of estimation that are related to income were noted in this table. These may be summarized as follows:

EXHIBIT 3.4: FREQUENCY OF DISCREPANCIES BETWEEN ESTIMATE AND IRS DATA PROVIDED BY APPLICANT

INCOME LEVEL AS REPORTED TO IRS AGI DISCREPANCY	Less than \$4,000	\$4,001 - \$9,999	\$10,000 or More	TOTAL
Overestimated	45 50.6 12.8	28 27.2 8.3	10 8 2.9	81 24.1
Accurate*	16 18.8 4.8	6 5.8 1.8	2 1.4 0.6	24 7.1
Underestimated	26 30.6 7.7	69 67. 20.5	136 91.9 40.5	231 68.8
TOTAL	85 25.3	103 30.7	148 44.	336 100.

* NOTE: For this table accuracy is defined as no difference between IRS-reported AGI and BEOG-reported AGI. (See text for explanation).

EXHIBIT 3.5: FREQUENCY OF AGI DISCREPANCIES BY FINAL AGI



- o 50.6 percent of low income applicants overestimated their 1974 income as compared to the IRS data. While only 18.8% of this category estimated accurately, these individuals accounted for 66.7% of all accurate estimators.
- o 91.9% of individuals with an actual AGI of \$10,000 or more underestimated their income. Levels of accurate and overestimating were correspondingly low. In the total sample only 68.8% underestimated at all, with approximately 51% underestimating by more than \$300.

In summary, this pattern exhibits a similar trend to the findings of the CRT Look-up Study comparison of discrepancies with income, where accuracy was strongly associated with low income. However, the degree of inaccuracy in the direction of underestimation at the high income level is substantially increased in this sample.

QUESTION #4: IS THE DISCREPANCY BETWEEN 1974-75 ESTIMATED INCOME AND 1974 FEDERAL INCOME TAX RETURN TAX RELATED TO THE DISCREPANCY BETWEEN 1975-76 REPORTED INCOME AND 1974 FEDERAL INCOME TAX RETURN?

In order to examine the 1975-76 reporting behavior of applicants who filed a supplemental in 1974-75, an analysis of the records of individuals selected for the Individual Case Follow-up was performed. One task in the Case Follow-up component of the Estimated Income Study was to search the 1975-76 BEOG master file to determine whether 1975-76 application had been submitted. This manual procedure used the microfiche copies of the master file records. These records were matched with the 1974 tax form supplied by the applicant for this investigation.

Records of the first 85 respondents who supplied documents containing AGIs that were not consistent with their 1975-76 BEOG AGI were studied. The consistency of reporting among the 1974-75 supplemental, the 1975-76 application, and the 1974 tax form was examined. Ideally, all three documents should match. It was not

expected that the 1974-75 supplemental would match either of the other two because the estimates are most often based on unknown quantities. However, the 1975-76 application should have had the same AGI reported as was entered on the IRS 1040, and the 1974-75 estimate should have been reasonably close to that figure. While investigating discrepancies in the CRT Look-up Study between the 1974-75 supplemental's estimate of 1974 income and the 1975-76 application's report of 1974 income, the question arose as to how accurate the report of income was. If the estimate was inaccurate compared to the IRS true income, and the report was also discrepant with the IRS data, to what extent are the inaccuracies being compounded or do they offset each other? Because the CRT Look-up Study did not involve follow-up of the applicants to obtain tax data, this question could only be answered with the present study component.

Comparisons of the discrepancies of the tax data with both the 1974-75 record and the 1975-76 record were made. Of the 85 in the sample, 34 individuals' records had discrepancies from tax data on their 1975-76 application which qualified them for larger grants than their estimated income had. That is, their discrepancies from true (IRS-reported) income, were, in most cases, larger for the year in which they were reporting actual income than they had been the previous year when they had estimated their income. Overall, the reports of income were discrepant by more than \$300 in 77% of the cases. This may be compared to the frequency of discrepancies found for the estimates which was 83 percent overall.

Six individuals had the same income reported on the 1974-75 supplemental application and the regular 1975-76 application. Four individuals from this group had significant discrepancies from true income (+\$2,252, +\$5,237, +\$12,729, and -\$11,103) indicating that they reported the 1974-75 estimate again in 1975-76, instead of reporting adjusted gross income from the Federal Income Tax return.

The remaining 39 individuals' change in discrepancy level had the effect of reducing their BEOG eligibility. Twelve persons who had large discrepancies on their 1974-75 estimates reported income in 1975-76 that was not significantly different from the IRS data. Several of these individuals however, decreased their eligibility from the amount their true income qualified them for by overreporting income to BEOGs.

QUESTION #5: ARE THE IRS FORMS SUBMITTED TO THE VALIDATION CONTRACTOR THE SAME AS THOSE SUBMITTED TO IRS FOR TAX PURPOSES?

The frequency and magnitude of discrepancies between the BEOG income estimate and the income contained in the IRS form submitted by the applicants in this study appear to indicate that true copies of the tax forms had been submitted. In addition, the distribution of discrepancies between the IRS forms and the 1974-75 estimated income appears to be similar to that of the discrepancies between the 1975-76 applications of BEOG recipients with their 1974-75 estimated income. Preliminary review of these findings would lead to a conclusion either of the validity or falsification of both the IRS and BEOG reported incomes provided to Basic Grants. However, this question may not be investigated in full without the results of the IRS component of the Estimated Income Study. Thus, a more thorough discussion will be presented in Chapter 4 dealing with the IRS component, and Chapter 5, overall Estimated Income Findings.

3.4: INDIVIDUAL CASE FOLLOW-UP CONCLUSIONS

The Individual Case Follow-up Study differed methodologically in several ways from the other components of the Estimated Income Studies. For one thing, the categories of discrepancy varied from the other two which were consistent with each other. Secondly, the analysis of discrepancies by income level used the IRS reported AGI rather than the standard estimate provided to BEOG. These two differences, however, enable the analysis of estimating accuracy to more fully explain the nature of errors in estimation.

The Individual Case Follow-up study found a higher rate of discrepancies greater than \$300 than did the CRT Look-up. This was caused by the use of a more stringent criterion document - the applicant's IRS return rather than the income reported to BEOG in the following year. Indeed, one of the analyses performed for the Individual Case Follow-up found that the 1975-76 application's report of 1974 income was significantly discrepant from both the 1974-75 estimate of income and the 1974 income tax return. About half of those discrepant on both years' applications were more discrepant from "true" income (IRS-reported) in 1975-76 than they had been on the 1974-75 supplemental.

The following table highlights similar findings between the CRT and Individual Validation studies:

	CRT Look-up Findings	Individual Validation Findings
Estimated income and Actual Income within:		
± \$200	24.7%	-
± \$300	28.5%	17.0%
± \$1,000	47.3%	36.3%
± \$2,000	64.7%	-
± \$3,000	-	67.5%

As is evidenced by this table, the Individual Validation Study found 11 percent fewer applicants to estimate accurately within \$300 than did the CRT Look-up study. This finding may be due to variations in the samples selected or in the study methodology. It may also indicate that people who estimate income in a prior year tend to report income accurately in the succeeding year. In both studies, it is important to note that more than 30 percent estimated inaccurately by greater than \$2,000 to \$3,000. This

implies that some income estimates are inaccurate by very large dollar amounts, and that the discrepancy ranges used in both studies were not large enough to provide a clear indication of the magnitude of discrepancies.

In the analysis of discrepancies by level of income in the Individual Case Follow-up component, the income data utilized were the AGIs found on the IRS 1040 or 1040A supplied by the applicants. While the other study components found that accuracy decreased as the level of the estimated income increased, the relationship between income and estimating accuracy was more dramatic when assessed against the applicants' true incomes. For example, 91.9 percent of individuals who had income levels over \$10,000 as reported to IRS underestimated their incomes. Another interesting comparison which will illustrate the finding was between the distribution of the sample for the estimated income levels and the IRS reported income levels:

INCOME RANGES	ESTIMATED INCOME	IRS REPORTED INCOME
Under \$4,000	25.5%	25.3%
\$4,000 - \$9,999	43.9%	30.7%
Over \$10,000	30.6%	44.0%

It would appear from the fact that many of the individuals sampled had large discrepancies between their 1975-76 applications and the IRS forms they provided to the validation contractor, that the individuals who provided the IRS data tended to provide "true" copies of their tax returns. If individuals had tended to supply false documentation for their IRS 1040s, then there would have been fewer discrepancies between the two AGIs. This conclusion was also upheld in the comparison of discrepancies with the IRS data base where it was found that the rates of discrepancy were similar

between the Individual Case Follow-up and the IRS sample study. Thus, almost all, if not all, of the individuals provided true copies of their IRS forms when asked to by Basic Grants.

Another difference between the Individual Case Follow-up and the CRT Look-up student components was the ability of the former to obtain data on the reasons supplemental applicants had for estimating their incomes. Applicants appear to use the separation/divorce criterion infrequently and tended not to respond to the Validation Contractor's requests for data as frequently as other applicants. Persons reporting death as the supplemental reason tended to be the most accurate, but again, accounted for a small portion of the sample. Overall, no strong relationship was found between the reason for filing supplemental applications and the extent of discrepancies.

A further discussion of the relationships among the three Estimated Income Components is presented in Chapter 5 to this report.

4

IRS SAMPLE STUDY

The third and largest component of the Estimated Income Study is the IRS Sample Study. Closely patterned after the IRS Sample Study for the regular Basic Grant applicant population,^{1/} the purpose of this component was to examine the overall accuracy of estimation, and to determine whether estimating accuracy was related to any applicant characteristics. The secondary purposes of this component were to compare its findings with those of the other two components and to examine the appropriateness of the measures used in other studies. This chapter discusses the objectives, methodology, findings and conclusions of the IRS Sample Study.

4.1: IRS SAMPLE STUDY OBJECTIVES AND ANALYSIS QUESTIONS

The IRS Sample Study component of the Estimated Income Study provides insight into each of the major study objectives and permits comparisons with the other study findings. Because all data and analyses were computerized and no individual contacts or manual data collection were necessary, a much larger sample was utilized for this study. However, available data elements were limited in number. This study examined how well individuals are able to estimate compared to the "true" incomes reported to IRS, obtained data on the

^{1/} See Volume III of this report for the IRS Sample Study for regular BEOG applicants

size, extent and nature of misreporting; and obtained evidence on the extent to which applicants, when asked to provide supporting documentation, submitted the copies of their tax forms. Specifically, the analysis questions to which this study component was addressed include the following:

- 1) What is the overall accuracy of estimation?
- 2) How is accuracy of estimation associated with applicant/program characteristics?
 - 2.1) Which categories of applicants reflect higher or lower discrepancy rates in estimated adjusted gross income (AGI)?
 - 2.2) Which categories of applicants reflect higher or lower discrepancy rates in estimated taxes paid?
 - 2.3) Which categories of applicants reflect higher or lower discrepancy rates in actual household size/exemptions?
- 3) Does the frequency of corrections made by applicants to their BEOG applications relate to misreporting?
 - 3.1) Does the frequency of corrections relate to the size of AGI discrepancies?
 - 3.2) Do applicants report accurate AGI's, then submit a supplemental application in order to increase the size of their awards?
- 4) Do applicants supply true copies of tax forms when participating in the validation effort?^{1/}
- 5) What are the effects of misreporting by the estimated income population on BEOG eligibility?

These analysis questions and subobjectives provide the framework for the IRS analysis and sampling methodology, to be described in the following sections.

^{1/} This analysis was performed only on the eligible population.

4.2: IRS SAMPLE STUDY METHODOLOGY

This section summarizes the sampling and analysis methodology of the IRS Sample component. Volume III to this report presents the IRS Sample Study procedures in greater detail. The Estimated Income Study IRS Sample component was arranged with the cooperation of both the Office of Education and the Internal Revenue Service. In drawing the sample and also in obtaining IRS cooperation, considerations were given to the types of data needed and to the need for protection of the privacy of individuals. Thus, all generated tables were computed on individual records, but displayed only for aggregates. The computer processing was handled by four different agencies, one of whom, including IRS, had access to the merged data file with identifiers.

4.2.1: IRS Sample Study Sample Selection

The population of primary interest in this study was comprised of applicants who had submitted a supplemental 1974-75 Basic Grant application,^{1/} and had received an eligibility index of 1200 or less, indicating that they were eligible to receive a Basic Grant award. It should be noted, however, that all applicants with eligible ratings do not necessarily receive awards because they may not enroll in an eligible institution. The 1974-75 Basic Grant data base contained records for 10,825 such applicants who had received eligibility indices of 1200 or less.

The second population of interest to this study consisted of 1974-75 applicants who had filed a supplemental application, but

^{1/} Applicants who suffer from specific circumstances which severely affect their financial status are allowed to file a supplemental application and report an estimate of the current year's income instead of the prior year's actual income.

received eligibility indices greater than 1200, indicating that they were ineligible to receive a Basic Grant. The Basic Grant data base included 6,535 such records.

Within the two populations of interest (eligibles and ineligibles), applicants were grouped according to three stratification variables hypothesized to be associated with discrepancies. The stratification variables selected were (1) financial dependency status, (2) income level, and (3) reason for filing supplemental. The following income levels of income were used:

1. For applicants dependent upon parental financial support, parents' adjusted gross income reported on the Basic Grant applications in the following ranges:
 - Less than \$4,000
 - \$4,000 to \$6,999
 - \$7,000 to \$9,999
 - \$10,000 to \$12,499
 - \$12,500 and over
2. For applicants independent of parental financial support, applicants' adjusted gross income reported on the Basic Grant application in the following ranges:
 - Less than \$4,000
 - \$4,000 to \$6,999
 - \$7,000 to \$9,999
 - \$10,000 and over

The third variable, reason for filing a supplemental application included four circumstances under which a supplemental in the 1974-75 application year could be submitted:

1. Death of parent/spouse
2. Unemployment of parent/spouse
3. Unable to pursue income-related activities
4. Divorce/separation of parent/applicant

The Office of Education specified that the total sample size contain approximately 6,000 1974-75 Basic Grant supplemental applicant records. An adjustment was made to this sample size in order to account for 20 to 30 percent of the dependents' whose parents' social security numbers were not provided in the application.^{1/} Thus, the sample size was increased to approximately 7,300. Sample sizes for each stratification cell were determined such that the sampling error was not greater than 5 percent at the 95 percent confidence interval. Exhibit 4.1 presents the strata counts and match rates for this study.

Sample selection was performed by sorting the application data base by income level and supplemental reason. Samples were then selected for each stratum, utilizing a random start and a skip interval equal to the proportion of the population to the desired sample size for each cell.

4.2.2: Confidentiality Measures

Effective security precautions were instituted in order to insure the privacy of individuals and the anonymity of both their Basic Grant and Internal Revenue Service Records. These are explained in detail in Volume III to this report. Briefly, they may be summarized as follows.

First, IRS data were not used as a basis for contacting an individual and/or changing that individual's Basic Grant record.

^{1/} The parent social security number was the primary means of linkage with IRS data for dependent application records. The 1974 Basic Grant applications requested, but did not require, the parents' social security numbers. In addition, only the applicant social security number is included in the automated data base; parent social security numbers remain only in the hard copy records. Thus, these data had to be extracted manually for this study.

**EXHIBIT 4.1: POPULATION, SAMPLE SIZES, AND MATCHED RECORDS FOR IRS SAMPLE STUDY
(ESTIMATED INCOME)**

ELIGIBLES

STRATA CODE	CLASSIFICATION (Eligibility, AGI, Dependency, Supplemental Reason)	UNIVERSE	SAMPLE	MATCHES
4111	Eligible, \$0-\$3,999, Dependent, Death	958	239	58
4112	Eligible, \$0-\$3,999, Dependent, Loss of Employment	1,241	310	225
4113	Eligible, \$0-\$3,999, Dependent, Unable to Pursue Income Activities	1,208	302	205
4114	Eligible, \$0-\$3,999, Dependent, Parents Separated/Divorced	85	85	31
4121	Eligible, \$0-\$3,999, Independent, Death	65	65	19
4122	Eligible, \$0-\$3,999, Independent, Loss of Employment	336	167	59
4123	Eligible, \$0-\$3,999, Independent, Unable to Pursue Income Activities	801	226	160
4125	Eligible, \$0-\$3,999, Independent, Applicant Separated/Divorced	167	167	64
	Subtotal	4,861	1,561	821
4211	Eligible, \$4,000-\$6,999, Dependent, Death	227	227	65
4212	Eligible, \$4,000-\$6,999, Dependent, Loss of Employment	1,093	273	220
4213	Eligible, \$4,000-\$6,999, Dependent, Unable to Pursue Income Activities	764	254	200
4214	Eligible, \$4,000-\$6,999, Dependent, Parents Separated/Divorced	90	90	49
4221	Eligible, \$4,000-\$6,999, Independent, Death	9	9	1
4222	Eligible, \$4,000-\$6,999, Independent, Loss of Employment	187	187	52
4223	Eligible, \$4,000-\$6,999, Independent, Unable to Pursue Income Activities	160	160	96
4225	Eligible, \$4,000-\$6,999, Independent, Applicant Separated/Divorced	13	13	10
	Subtotal	2,543	1,213	693
4311	Eligible, \$7,000-\$9,999, Dependent, Death	112	112	29
4312	Eligible, \$7,000-\$9,999, Dependent, Loss of Employment	1,109	227	235
4313	Eligible, \$7,000-\$9,999, Dependent, Unable to Pursue Income Activities	612	203	165
4314	Eligible, \$7,000-\$9,999, Dependent, Parents Separated/Divorced	48	48	32
4321	Eligible, \$7,000-\$9,999, Independent, Death	4	4	1
4322	Eligible, \$7,000-\$9,999, Independent, Loss of Employment	46	46	15
4323	Eligible, \$7,000-\$9,999, Independent, Unable to Pursue Income Activities	49	49	22
4325	Eligible, \$7,000-\$9,999, Independent, Applicant Separated/Divorced	0	—	—
	Subtotal	1,980	689	499
4411	Eligible, \$10,000-\$12,499, Dependent, Death	41	41	11
4412	Eligible, \$10,000-\$12,499, Dependent, Loss of Employment	641	319	250
4413	Eligible, \$10,000-\$12,499, Dependent, Unable to Pursue Income Activities	315	315	251

EXHIBIT 4.1: (Continued)

4414	Eligible, \$10,000-\$12,499, Dependent, Parents Separated/Divorced	13	13	6	
4421	Eligible, \$10,000-\$12,499, Independent, Death	0			
4422	Eligible, \$10,000-\$12,499, Independent, Loss of Employment	3	3	0	
4423	Eligible, \$10,000-\$12,499, Independent, Unable to Pursue Income Activities	1	1	0	
4425	Eligible, \$10,000-\$12,499, Independent, Applicant Separated/Divorced	0			
	Subtotal	1,014	692	518	
4511	Eligible, \$12,500 and over, Dependent, Death	16	16	6	
4512	Eligible, \$12,500 and over, Dependent, Loss of Employment	297	297	225	
4513	Eligible, \$12,500 and over, Dependent, Unable to Pursue Income Activities	107	107	79	
4514	Eligible, \$12,500 and over, Dependent, Parents Separated/Divorced	6	6	5	
4521	Eligible, \$12,500 and over, Independent, Death	0			
4522	Eligible, \$12,500 and over, Independent, Loss of Employment	0			
4523	Eligible, \$12,500 and over, Independent, Unable to Pursue Income Activities	1	1	0	
4525	Eligible, \$12,500 and over, Independent, Applicant Separated/Divorced	0			
	Subtotal	427	427	315	
	ELIGIBLE	TOTAL	10,825	4,582	2,846

INELIGIBLES

STRATA CODE	CLASSIFICATION	UNIVERSE	SAMPLE	MATCHES	
3110	Non-Eligible, \$0-\$3,000, Dependent	1,179	383	211	
3120	Non-Eligible, \$0-\$3,000, Independent	276	276	160	
3210	Non-Eligible, \$4,000-\$6,999, Dependent	941	313	201	
3220	Non-Eligible, \$4,000-\$6,999, Independent	239	239	150	
3310	Non-Eligible, \$7,000-\$9,999, Dependent	1,055	351	243	
3320	Non-Eligible, \$7,000-\$9,999, Independent	170	170	85	
3410	Non-Eligible, \$10,000-\$12,499, Dependent	840	419	318	
3420	Non-Eligible, \$10,000-\$12,499, Independent	77	77	37	
3510	Non-Eligible, \$12,500 and over, Dependent	1,712	428	333	
3520	Non-Eligible, \$12,500 and over, Independent	46	46	18	
	INELIGIBLE	TOTAL	6,555	2,712	1,756

Rather, the purpose of this study was to determine the nature and extent of Basic Grant misreporting in the population, so that changes could be made to the application processing system in the future to avoid program abuse.

Secondly, although results could be presented in terms of groups/categories of applicants, analysis required the matching of individual records for comparisons. Thus, rigid precautions were instituted to assure that individually identifiable IRS data could not be accessed by OE and that individually identifiable BEOG data could not be accessed by IRS. No agency involved in the processing possessed individually identifiable records of both BEOG and IRS data, and all intermediate files were destroyed.

In addition, data were suppressed in all analysis table cells which included less than ten (10) observations.

4.2.3: Analysis Methodology

In order to develop its analysis file, IRS attempted to match all social security numbers and last names contained on the Cross-Reference file with the IRS Master File, which contains a record of all persons filing a Federal Income Tax return. A complete "match" was defined as a match on social security number and last name, and the presence of a 1974 return. IRS was requested to supply data for instances where a match was found on the social security number, but not on the last name. The file generated by IRS contained the 1974 AGI, total acceptable tax, exemptions, filing date, and filing status. The BEOG and IRS data were then merged (see Volume III, Appendix A).

Cross-tabulations were generated for each analysis objective utilizing the Applied Management Sciences' PROTABS generalized table producing package. Three sets of tables were created: for eligibles, for ineligibles, and for the recalculation of eligibility indices.

There tables, produced for the sample initially, were inflated for the total study group using the appropriate inflation factors for analysis.

The tables were inflated to represent that portion of the population which was based on known data. That is, if an individual's record did not have a match on the IRS file, that record was not considered in inflation. Thus, although the population of the eligible applicants total population of 10,825 individuals, the tables generated approximately 7050 matched individuals; the remainder were non-matches. Tables were based on a sample from the 7050, and inflation factors were utilized to approximate this total. Weighting of the data was performed in order to enable comparisons among the data for representatives of all sampling strata. In order to minimize sampling error, different sampling proportions were appropriate for the various strata. For example, a stratum with few cases is often sampled with certainty (a proportion of 1) because the elimination of even one member will cause the error to be larger than that established for the needs of the study. A stratum with many hundreds of members, however, may be sampled with a much smaller proportion, because beyond a certain point adding cases or increasing the proportion by any degree will not yield an appreciable lower rate of error. Because the sample frame for the estimated income study had many cells, and several different sampling proportions were used, weighting of the cells to represent the population sampled was performed to enable statistical analysis across the strata of the population. Because of the high rate of non-match, and because only a few characteristics of the non-matched group could be identified, all analyses were performed on the study group, inflated to represent the population based on the matched group, without accounting for the non-matched group. Therefore, this report contains findings for the matched study group and not for the total 1974-75 supplemental population.

4.2.4: Specification of Variables

Because of the potentially large number of variables available, careful attention was paid to the specification of the dependent and independent variables. The first consideration was given to those variables which appeared to have the most bearing on the study objectives. Secondly, the comparability of definitions of data items between the Basic Grant application and IRS forms 1040 and 1040A was evaluated. A review of the 1974-75 Basic Grant application and the 1974 IRS forms 1040 and 1040A revealed that one data item, definitionally constant on both agencies' forms, was important to analysis.

1. Adjusted Gross Income (AGI)

Applicants who file a supplemental application and estimate their income do not report the adjusted gross income earned in the year previous to the application. Rather, some circumstances have occurred causing their current year's income to be substantially changed from the previous year's. Thus, they are asked to estimate for the current year. The definition used by Basic Grants for adjusted gross income is identical to that used by IRS. The estimate is calculated by the applicant by adding the AGI already received by the date of the application to an estimate of the amount expected for that part of the year that has not yet occurred. Because the IRS-BEOG comparison centers around the estimate of 1974 income, the IRS component of the Estimated Income Study concentrated on comparisons of the supplemental application with the 1974 tax return.

Definitions of the other data items were found to be similar but not exactly the same between the agency forms. These are:

2. Taxes paid, and
3. Household size.

The taxes paid variable is calculated by Basic Grants for applicants who estimate their income based on the formula used by IRS.

The IRS Master File, to which these data are comparable, however, does not have a field directly applicable to "taxes paid"; it does have "total acceptable tax" which includes taxes paid, and, for approximately five percent of taxpayers, such items as self-employment tax and tax credits. This was considered to represent a small amount of error. Thus, the taxes paid variable was analyzed.

The household size variable is similar to the exemption variable utilized by IRS. However, because of confounding definitional problems, only one exemption was counted for the primary and secondary taxpayers claiming additional exemptions for age and/or blindness, and, for analysis, a difference of one (1) between the BEOG household size and IRS exemption fields would account for definitional differences. Therefore, differences of +1 were considered to be insignificant.

In addition to the above comparison variables, a fourth dependent variable was examined. This was the eligibility index as derived from IRS data as opposed to Basic Grant-reported data. Because of the complexity of the BEOG eligibility formula it was of prime importance to study the effects of discrepancies in all comparative variables on the student's awards. Some discrepancies in the application might offset one another. More importantly, each variable enters into the formula in a unique way and may have disproportionate effects on the outcome of the eligibility calculation. Only through the examination of the differences between the original BEOG eligibility index and the IRS data could an estimate be made of the impact on payments that may have been granted to applicants because financial information had been misreported or poorly estimated. From this analysis, revised procedures may be indicated for the estimating process.

Applicant categories comprise the independent variables for this study. These variables were chosen which were found, through related research and preliminary results from tables generated, to have

possible relationships with discrepancies. Those independent variables eventually selected include:

- Eligibility
- Dependency Status
- Adjusted Gross Income Level
- Reason for filing supplemental application
- Filing Status - applicants/parents reporting divorced/separated/widowed to BEOG and filing either a joint or separate IRS return
- Date of signature of BEOG applications
- Applicants who had changed adjusted gross income on the Basic Grant application from correct to incorrect
- Applicants making changes in their Basic Grant application.

The dependent and independent variables listed above were the means through which the analysis questions were evaluated. Section 4.3 presents the findings related to the analysis questions. Section 4.4 will discuss conclusions drawn from these findings.

4.3: IRS SAMPLE STUDY FINDINGS

Comparative data for the 1974-75 estimated income population with the 1974 income tax return were available for 2,840 eligible applicants and 1,750 ineligibles. Because of the high rate of non-matches, cross-tabulations were performed with percentages based on study group totals. This method used the total "matches" as the basis upon which percentages were calculated. Any individual for whom there was not a matched record in the IRS and Basic Grant Master Files was not counted. For all tables, data suppressed for confidentiality purposes were treated as "zeroes" or non-matches. These data are presented by asterisks (*). The text below presents summary tables as necessary.

The rate of non-matches, which was 35 percent overall for the estimated income population, had serious implications for any conclusions to be drawn. On the one hand, discrepancy and accuracy

rates seem inordinately low if one-third of the population cannot be put into one of the discrepancy categories. On the other hand, small numbers of non-matches might be assumed to have been distributed as was the matched group. Even if they are not, the error introduced by such an assumption of equality is small. However, with large groups of unknowns, similarity of distribution cannot be assumed. The error introduced can be quite large. Indeed, the groups with the lowest percentage-of-total-population discrepancies often had high non-match rates, and when the percentage-of-matched-group discrepancy rate was calculated, it was found to be quite high. For this reason, only percentages of the known population, or "study group" are discussed in this report.

Two groups which had lowest match rates have a significant impact on the study findings because they represent two of the major groups of interest -- independents and low income dependents. There are two hypotheses which may explain the low match rate for independents. The first is a function of the IRS master file. This file contains only one social security number per record. Consequently, the secondary taxpayer filing a joint return is not identified on the master file. This implies that it is likely that many of the non-matched independent records were those of married applicants who filed a joint return as the secondary taxpayer. The second explanation regarding the high non-match rates for independents could be that many "independents" do not file tax returns either because they are really dependent and are claimed on their parent's tax return as exemptions, or because their income is quite low and they are not required to file.

The second large applicant group which had a low match rate was applicants with incomes under \$4,000. A low match rate for this group was anticipated because many persons in this income range do not have an earned income large enough to necessitate filing a Federal Income Tax return. The loss of this group from the study

findings, unfortunately, does cause an obvious bias -- because IRS records are not likely to be available for a large portion of "true" low income people, it can be expected that the matched set of records used in the analysis may be comprised largely of persons who reported a low income to BEOG, but reported a high income to IRS.

A complete discussion of the effects of non-match on the validation study findings may be found in Volume III, Chapter 2. The non-match rate for the estimated income population tended to be lower than that for the general population.

4.3.1: Characteristics of the Matched Sample

Because the stratified random sampling procedure utilized in this study involved nonproportional sampling, data obtained were inflated by using sample proportions to represent the Basic Grant applicant population.^{1/} The characteristics of the population and matched study group are presented in Exhibit 4.2 to facilitate interpretation and evaluation of the study findings. This exhibit displays the portions of the population and sample that are comprised of applicants with characteristics on specific levels of the sampling variables.

As the table shows, the study groups for both eligibles and ineligibles are skewed in terms of the distribution of dependents and independents. The eligible study group had fewer independents than the population, while the ineligible study group had more. Also, in both study groups dependents had fewer individuals in the under \$4,000 income category than did the populations, while independents had more individuals in this category for both populations. In addition,

^{1/} Inflation factors were based on the sample proportion relative to the population. Because these were often not whole numbers, totals may vary from table to table due to rounding errors.

EXHIBIT 4.2: IRS STUDY - CHARACTERISTICS OF THE POPULATION AND STUDY GROUP

ELIGIBLES		
VARIABLE	POPULATION %	STUDY GROUP %
<u>Dependency Status</u>		
Dependent	83.0	87.8
Independent	17.0	12.2
<u>Reason for Filing Supplemental</u>		
Death	13.2	4.2
Loss of Employment	45.8	52.8
Unable to Pursue Income Activities	37.1	41.4
Separation or Divorce	3.9	1.6
<u>Income Level - Dependent</u>		
Under \$4,000	38.9	32.7
\$ 4,000 - \$ 6,999	24.2	26.3
\$ 7,000 - \$ 9,999	20.9	24.9
\$10,000 - \$12,499	11.2	12.3
\$12,500 and over	4.7	3.8
<u>Income Level - Independent</u>		
Under \$4,000	71.9	91.2
\$ 4,000 - \$ 6,999	19.4	8.8
\$ 7,000 - \$ 9,999	7.7	*
\$10,000 - \$12,499	0.9	*
\$12,500 and over		

EXHIBIT 4.2: (Continued)

INELIGIBLES		
VARIABLE	POPULATION %	STUDY GROUP %
<u>Dependency Status</u>		
Dependent	87.6	81.2
Independent	12.4	18.8
<u>Reason for Filing Supplemental</u>		
Death	25.9	14.1
Loss of Employment	41.2	49.7
Unable to Pursue Income Activities	31.5	35.4
Separation or Divorce	1.4	0.3
<u>Income Level - Dependent</u>		
Under \$4,000	20.6	15.1
\$ 4,000 - \$ 6,999	16.4	15.2
\$ 7,000 - \$ 9,999	18.4	18.5
\$10,000 - \$12,499	14.7	16.3
\$12,500 and over	29.9	34.3
<u>Income Level - Independent</u>		
Under \$4,000	34.2	41.3
\$ 4,000 - \$ 6,999	29.6	39.4
\$ 7,000 - \$ 9,999	21.0	13.1
\$10,000 and over	15.2	5.6

applicants who reported death as the reason for filing a supplemental also evidenced a low match rate since the deceased person obviously would not have filed an IRS return. These changes in distribution have significant impacts on findings, particularly for the independents, for whom the non-match rate was extremely high. Findings for the independents, therefore do not have a high degree of reliability.

4.3.2: Findings Related to the Analysis Questions

This section presents the findings of cross-tabulations performed to answer the analysis questions listed in Section 4.1. The major purpose of the Estimated Income IRS Sample Study was to obtain information related to the size and extent of the difference between information provided on the Basic Grant supplemental application and that reported to the Internal Revenue Service on the IRS 1040 and 1040A forms. The research questions posed in this study examined the nature of estimation errors in considerable detail utilizing a large number of variables. The findings are presented in three parts. Section 4.3.2.1 presents a summary of overall findings for the IRS Sample Study. Section 4.3.2.2 discusses the eligible population, and Section 4.3.2.3 presents the findings related to ineligible Grant applications.

4.3.2.1: Overall Findings Obtained from IRS Sample Study

In the 1974-75 BEOG application year, approximately 17,000 individuals used the supplemental application process. Of these, 62.4 percent were found to be eligible to receive a grant based on the estimate of income. In the regular application process, the Basic Grant program received 1,090,000 complete applications, and, of these, 61.2 percent had eligibility indices of less than 1200, indicating eligibility. The difference in eligibility between the estimated and the regular applications was not significant, although it might have been assumed that supplementals would have a higher rate of eligibility due to the nature of the application regulations. It might

be expected, for example, that persons who meet the criteria for filing estimated income would be restricted to those cases where income was likely to drop significantly, and hence, would result in a higher frequency of eligibility.

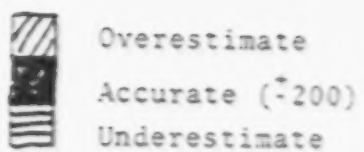
Because the estimated income was based, in most cases, on incomplete knowledge of the year's income, and would be compiled with the emotional trauma associated with some natural disaster, it was expected that the accuracy of estimation would be quite low. Tables generated for the IRS Sample component of the Estimated Income Study showed, in fact, that 13.6 percent of the study group estimated their adjusted gross income (AGI) accurately to within [±] \$200 of their actual income for 1974. Accuracy was less frequent in the IRS Sample than was true of the samples obtained for other Estimated Income studies. However, the criterion of accuracy was more restrictive^{1/} in the IRS comparison than it had been in the other study components. Exhibit 4.3 illustrates the overall accuracy of the Estimated Income IRS Sample.

The analyses of adjusted gross income discrepancies based on the IRS sample also showed that 68.5 percent of the sample underestimated their income on the 1974-75 supplemental. More than half of those who were significantly discrepant had discrepancies from true income of more than \$2,000. The ineligibles had somewhat fewer underestimators (65.5 percent) than did the eligibles (70.3 percent).

Overestimators represented 17.3 percent of the total study group, with a smaller portion of eligibles being in this category than

^{1/} For example, in the Individual Validation study, children of divorced parents could indicate the portion of adjusted gross income provided by the parent supporting the applicant. In the IRS study such determinations could not be made.

EXHIBIT 4.3 OVERALL ACCURACY OF ESTIMATED INCOME POPULATION



ineligibles (16.6 and 20.0 percent, respectively).

In general the ineligible had higher rates of accuracy and overestimating than did the eligibles, and lower rates of underestimating. This was also true when the dependent and independent populations were examined separately. The table below summarizes this finding:

Discrepancy Type	DEPENDENTS			INDEPENDENTS		
	Ineligible	Eligible	Total	Ineligible	Eligible	Total
Underestimate	66.4%	71.2%	69.3%	56.2%	63.8%	61.4%
Accurate (± 200)	12.7	11.2	11.8	32.3	26.8	28.6
Overestimate	20.9	17.6	18.9	11.5	9.3	10.0

As may be seen, the eligibles among both groups exhibited more underestimating than the ineligible. Independents appeared to have an accuracy rate more than twice as high as the dependents. Overall, 28.6 percent of independents were accurate estimators while only 11.2 percent of dependents were accurate.

Dependents, however displayed a somewhat higher tendency to overestimate their income. Almost 19 percent of the eligible dependents overestimated, compared to only 10 percent of the independents. Among dependent ineligible 20.9 percent overestimated their income. Overestimation generally decreases the applicants' eligibility for a Basic Grant, since it can raise the eligibility index.

For dependent eligibles, the frequency of accuracy decreased as the income level of applicants increased, while for dependent ineligible the opposite was true. For eligibles, 13.4 percent of those dependents with estimated incomes^{2/} under \$4,000 were accurate, and only 3.7 percent of those with estimated incomes over \$12,500 were accurate.

^{2/} In the IBS Study, the income estimated to BEC was used as the independent variable.

However, for ineligible, 7.5 percent in the lowest income category were accurate while 14.5 percent of those in the highest income category were accurate. An opposite trend was evident for underestimating. Here the highest frequency of discrepancies occurred in the lowest income category for ineligible and in the highest for the eligible population. Overestimating, however, did not appear to be related to income level.

Persons filing a supplemental application because they or their main source of support in their family were unable to pursue income activities for more than ten successive weeks in 1974 were the most accurate and the most likely to overestimate their incomes of the study groups classified by supplemental reason. Although underestimators in this category accounted for over 39 percent of the study group as a whole, only 64.8 percent of the people in this category underestimated by more than \$200. Persons citing death and separation or divorce were most likely to underestimate significantly; however, these categories accounted for the smallest proportion of the study group, with only 9.0 percent of the study group in these two categories combined. Those citing loss of employment had a 70.3 percent rate of underestimating, somewhat higher than the unable to pursue income activities category.

Overall, 12.4 percent of the study group was accurate with the frequency across categories ranging from 7.9 percent for those citing death to 16.0 percent for those citing unable to pursue income activities. In general, accuracy rates were higher for ineligible than for eligible, although, in the unable to pursue income activities category, the rates of accuracy were the same for both populations. Overestimating was more frequent for ineligible than for eligible (20.3 and 16.3 percent, respectively) and was especially high for the ineligible who cited unable to pursue income activities (24.4 percent). This group also had the lowest rate of overestimating (19.3 percent), compared to 67.4 percent for the eligibles in that category and higher rates of overestimating for other categories. Exhibit 4.4 displays a summary of reporting discrepancy rates by eligibility status and supplemental reasons.

EXHIBIT 4.4: FREQUENCY OF DISCREPANCIES BY SUPPLEMENTAL
REASON - TOTAL STUDY GROUP

AGI DISCREPANCY TYPE	SUPPLEMENTAL REASON				
	Total	Death	Loss of Employment	Unable to Pursue Income Activities	Separation or Divorce
Total Study Group	11,155 100.0	890 7.9	5,757 51.6	4,383 39.3	125 1.1
Total Positive Dis- crepancies Greater than +200	7,696 68.3	698 6.3 78.4	4,048 36.3 70.3	2,838 25.4 64.8	112 1.0 89.6
Total Accurate (-200)	1,450 12.4	70 0.6 7.9	664 5.4 11.5	703 6.3 16.0	13 0.1 10.4
Total Negative Dis- crepancies Less than -200	2,009 18.0	122 1.1 13.7	1,045 9.4 18.2	842 7.5 19.2	*

Count
Total%
Column%

The number of discrepancies found for adjusted gross income was the largest of any analysis. Taxes paid showed many fewer discrepancies than AGI, possibly because of the small levels of income found for the majority of the population. For example, 58.6 percent of the eligibles had no significant discrepancies. For ineligibles 46.4 percent had accurate taxes paid. In the household size/exemptions analysis, over 90 percent of the study group had reported accurately (+1) to BEOG.

The remaining analysis, the accuracy of the eligibility index, displayed an increase in accuracy over the AGI analysis as well. It appeared that, although most of the study group had discrepant AGIs, many of the discrepancies had little effect on the applicants' eligibility for a Grant. For example, over 50 percent of the eligible applicants had accurate SEIs (within 200 points) while only 13.1 percent had accurate AGIs (within \$200). The analysis of SEI accuracy was performed by substituting IRS data for AGI, Taxes Paid and Household Size for the BEOG-reported data elements. The accuracy of ineligibles also increased for SEI, although not as much. Further, for the ineligibles, only 5 percent of the applicants would have been eligible for a Grant based on their IRS reported data.

For the total study group, approximately 17 percent would have had their eligibility status changed either to or from eligible on the basis of IRS data. Overall SEI accuracy rates for the estimated income study group are presented in Exhibit 4.3. The comparison of SEI discrepancies between the eligibles, ineligibles and total study group is presented in Exhibit 4.6.

This section has presented an overall summary of the findings of the Estimated Income Studies IRS Sample Component, including comparisons of major subgroups. The remainder of this chapter will include specific study findings and conclusions based on these findings. The discussion of findings follows this section. It is divided into separate portions for the eligible and ineligible populations. Conclusions are presented in section 4.4 of this chapter.

EXHIBIT 4.5: OVERALL ACCURACY OF STUDENT ELIGIBILITY INDICES
BASED ON AGI PLUS TAXES PAID SUBSTITUTIONS -- TOTAL
STUDY GROUP

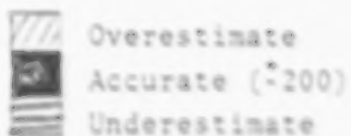
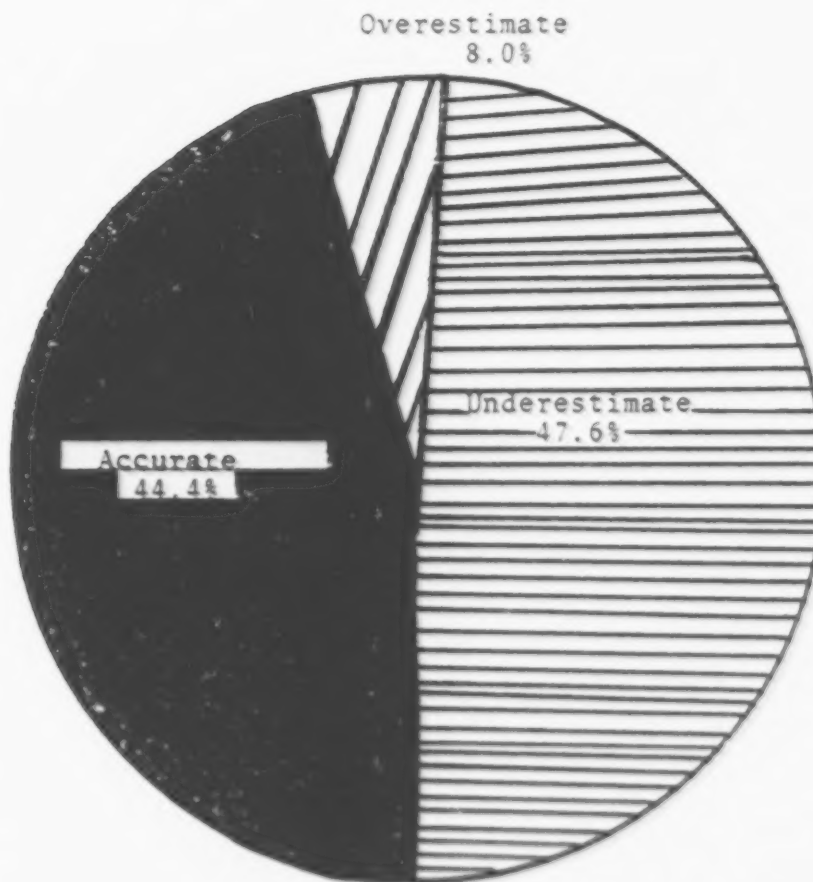


EXHIBIT 4.6: FREQUENCY OF DISCREPANCIES IN STUDENT
ELIGIBILITY INDICES - TOTAL STUDY GROUP

SEI DISCREPANCY TYPE	TOTAL	ELIGIBLES	INELIGIBLES
Total Study Group	11,344 100.0	7,031 62.0	4,313 38.0
Total Positive Discrepancies Greater than +200	5,400 47.6	3,059 26.9 43.5	2,341 20.6 54.3
Total Accurate (+200)	5,042 44.4	3,629 31.9 51.6	1,413 12.5 32.8
Total Negative Discrepancies Less than -200	902 8.0	343 3.0 4.9	559 4.9 12.9

Count
Total %
Column%

4.3.2.2: Findings Obtained for the Eligibility Study Group

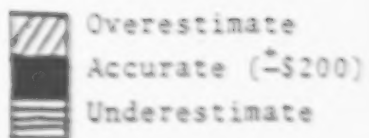
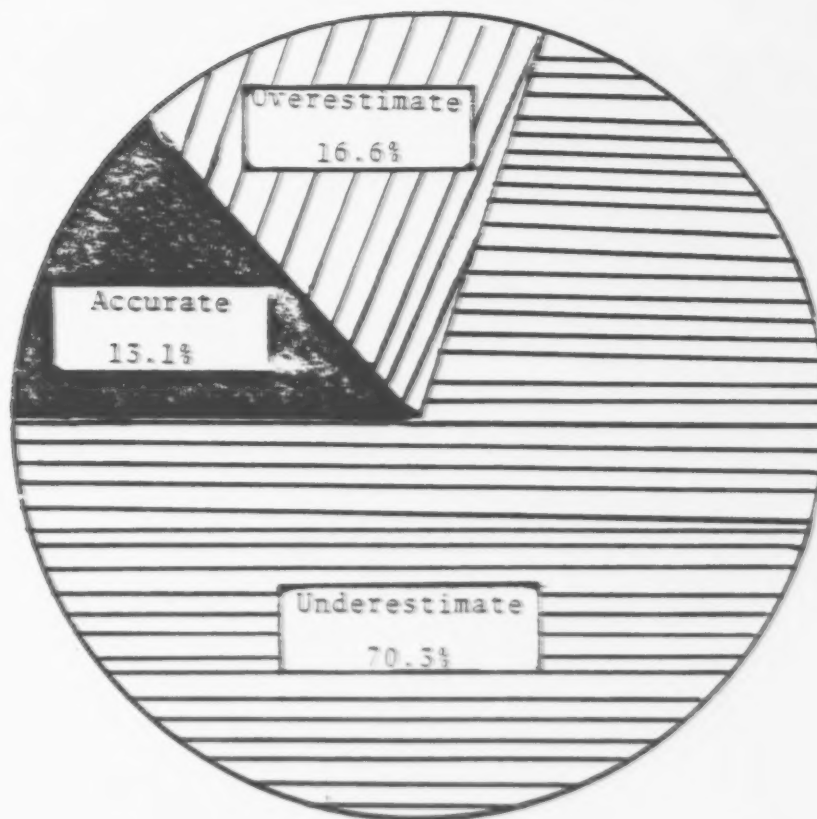
The eligibility of a Basic Grant applicant is determined by a complex calculation of the student's or his/her parents' ability to support his/her postsecondary educational pursuits. Several factors in varying degrees of importance are considered, including adjusted gross income, taxes paid, household size/exemptions, in the formula. Applicants receiving an eligibility index of 1200 or lower are qualified to receive a Grant. The lower the index, the higher the maximum level of the Grant (up to \$1400). This section of the report discusses the accuracy of estimating by those individuals who were eligible, using the estimate of income in the formula, to receive a Grant. Ineligibles are discussed in Section 4.3.2.3. Both this section and the section on ineligibles are arranged in the order of the analysis questions listed above in Section 4.1.

QUESTION #1: WHAT IS THE OVERALL ACCURACY OF ESTIMATION?

Approximately 64.8 percent of the eligible population had BEOG-IRS matched data. The dependent eligibles had a higher match rate than independents. Findings are presented in terms of the distributions of discrepancy types in the known or matched population. Therefore, it is not appropriate to generalize these findings to the total BEOG population.

The study group included an estimated 7,050 individuals. Of this total, 86.9 percent had estimated incomes which were discrepant with IRS data. Only 13.1 percent (see Exhibit 4.7) of the eligibles studied were able to correctly estimate 1974 income on the 1974-75 applications (within a tolerance of $\pm$ \$200). The majority of discrepancies, accounting for 70.3 percent of the eligibles as a whole, were positive, meaning that these individuals estimated that they would receive less income in 1974 than they actually did receive. Approximately half of all eligible underestimators had estimates which were more than \$2,000 different from their eventual income

EXHIBIT 4.7: OVERALL ACCURACY -- ELIGIBLE STUDY GROUP

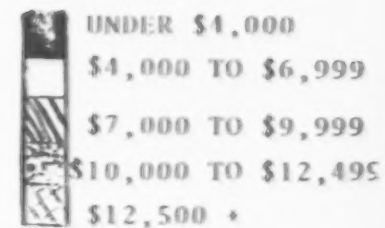


for 1974 (as later reported to IRS). In addition to the underestimators, 16.6 percent of all eligibles overestimated their income on their supplementals. That is, they thought that they would earn more in 1974 than they finally did. Persons who underestimate generally receive a lower eligibility index and a larger grant than they deserve, according to their "true" income. Persons who overestimate may be getting a smaller grant than they are entitled to.

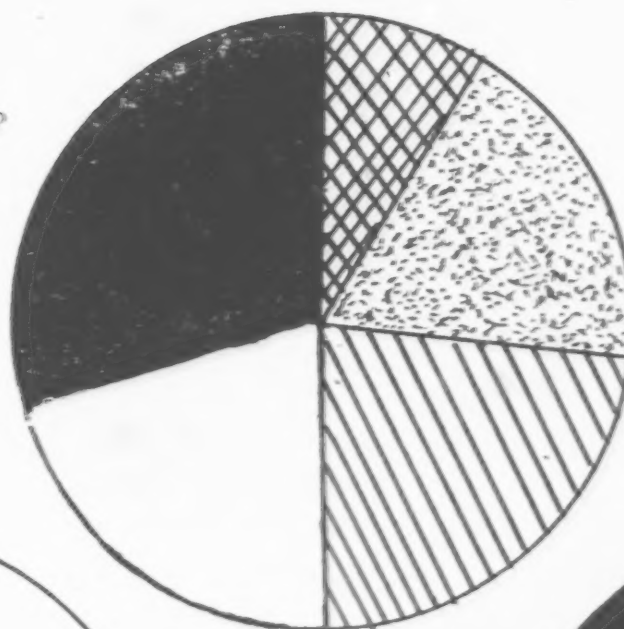
It would appear from the results summarized above that 86.9 percent of the study group with eligibility indices qualifying them for a grant did not receive an equitable index when estimated income is compared with true income. However, Question #5 in this section investigates the effects of the adjusted gross income on the eligibility index: over 50 percent of the eligibles were found to have insignificant discrepancies in their eligibility index after "true" AGIs were substituted in the eligibility formula for the discrepant amounts reported to BEOGs. The size of negative discrepancies (resulting from overestimates of AGI) was also smaller in the eligibility index calculations. However, most of the positive discrepancies in eligibility indexes (underestimates of income) remained in the largest discrepancy range.

The effects of recalculation of eligibility indices may be illustrated through the analysis of the way in which the change in index would have changed eligibility levels. Exhibit 4.8 presents these results. Almost 25 percent of the eligible study group would have been ineligible to receive any sized grant if their eligibility had been based on the adjusted gross income reported to IRS rather than their estimate. The effects of AGI discrepancies were minimized, however, by the fact that for the estimated income population, taxes paid, which is a factor in index calculations, was computed by Basic Grants rather than the applicants. Thus, taxes paid figures were discrepant because the AGIs on which they were based were discrepant, but there was no potential for misreporting or

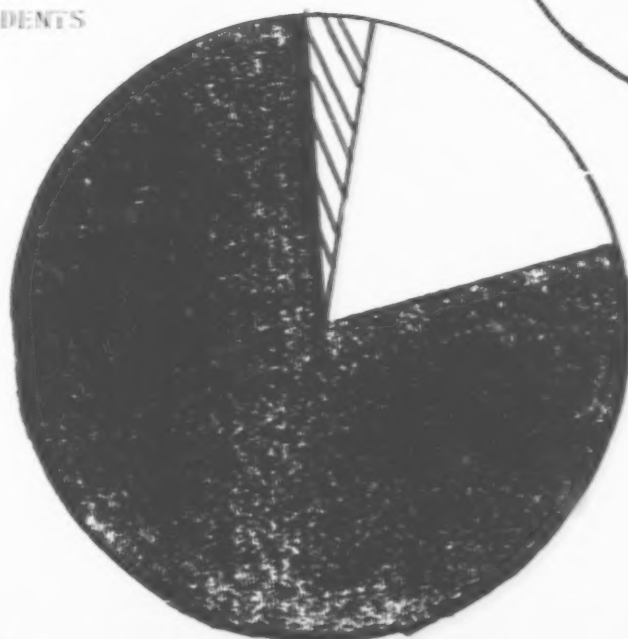
EXHIBIT 4.8: DISTRIBUTION OF INCOME LEVELS FOR IRS -- INELIGIBLE PORTIONS OF THE BEOG STUDY GROUP



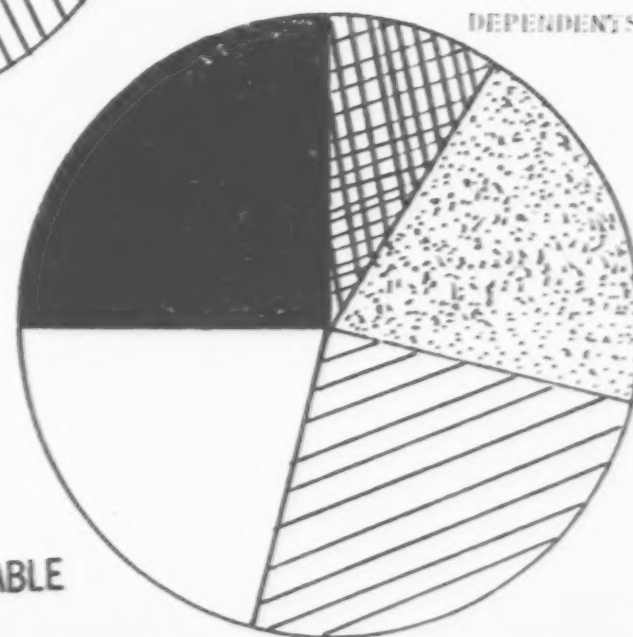
TOTAL STUDY GROUP



INDEPENDENTS



DEPENDENTS



BEST COPY AVAILABLE

poor estimating in the taxes paid figure. Substituting the IRS data on taxes paid and AGI for the corresponding data in the BEOG master file results in a total of 23.2 percent of the study group who would have been ineligible if their index had been calculated using IRS data.

The eligible estimated income population, then, has a much higher discrepancy rate than those persons filing regular applications, but the effects of discrepancies in adjusted gross income, especially, are minimized by the eligibility index formula and by the fact that supplemental applicants' taxes paid is calculated by Basic Grants. (See Question #5 for a discussion of SEI discrepancies.) Thus, as few as 50.7 percent of the eligible estimated income were underestimating to such a degree that their Grant eligibility was significantly lower than it would have been if they had been able to accurately estimate. The 50 percent figure is considered relatively small because AGI discrepancies had been found to occur for more than 70 percent of all income estimators.

It should be remembered at this point that the Basic Grant regulations were designed to allow estimation by as small a portion of applicants as possible precisely because estimators are not expected to be accurate. For instance, it is likely that a large portion of the estimations more accurately reflected the study group's 1974 income than the report of 1973 income would have. Because of the circumstances surrounding the applicant who meets one of the criteria for filing a supplemental, the estimation procedure enables these individuals to receive necessary assistance at the time it is needed. Generally, then, with a large portion expected to be discrepant, it can be seen that for half of the estimators, the estimation process appears to work. Only 50 percent may be considered to have had estimates poor enough to significantly affect the size of their Grant award. This may be further

evaluated in light of the fact that for the total population of persons eligible to receive a Basic Grant in 1974, those whose poor estimation significantly affected their eligibility accounted for only 0.5 percent of the total persons eligible to receive a grant. Thus, those who received an eligibility index of less than 1200 but who would have been ineligible based on data supplied to IRS amounted to approximately 0.2 percent of the total eligible BEOG population.

QUESTION #2: HOW IS ACCURACY OF ESTIMATION ASSOCIATED WITH APPLICANT/PROGRAM CHARACTERISTICS?

The analyses performed to investigate this question represented a complex series of cross-tabulations involving the major independent variables. The criteria of accuracy studied were:

- Adjusted Gross Income discrepancies between BEOG estimate and IRS reports of income;
- Taxes Paid discrepancies between BEOG supplemental application and IRS data; and
- Household size/exemptions discrepancies.

Question #2.1: Which categories of applicants reflect higher or lower discrepancy rates in adjusted gross income?

The discrepancies between BEOG estimates and IRS reports of applicant/parent income for 1974 were cross-tabulated with a series of applicant characteristics such as dependency status, income levels and supplemental reason. The results of each analysis for the eligible estimated income study group are presented below.

Dependency Status and AGI

For adjusted gross income discrepancies, dependency status was investigated by itself and with adjusted gross income. Dependents and Independents are considered quite different from each other in many of their financial characteristics as well. For example, in the BEOG eligibility calculations, and in the 1974-75 BEOG application year, independents who became unemployed could not use this situation as a reason for filing a supplemental application. Dependents whose parent(s) suffered a loss of employment for more than ten consecutive weeks could estimate their income using this

situation as the criterion for the supplemental. Independents also were required to have an extremely low income in order to be eligible while the income for dependents may have been considerably higher, since dependents frequently have more income offsets than independents. Because of this, the analysis of income characteristics relative to discrepancy ranges was performed separately for the dependents and independents who estimated income.

On the whole, independents appeared to be somewhat more accurate than dependents. In Exhibit 4.9 it can be seen that the rate of accuracy for independents was 26.8 percent compared to the rate for dependents which was 11.2 percent. The higher accuracy rate for independents may be due to several factors:

- 91.2 percent of the independent study group had estimated their incomes to be less than \$4,000, while only 32.7 percent of dependents were in this category
- 53.5 percent of the independent population could not be matched with IRS data, probably due to low income levels (meaning that they may not have been required to pay taxes), to the misreporting of dependency status, or because married students may have filed joint returns with their spouse.
- Independents comprised only 12.2 percent of the study group (and only 17 percent of the 1974-75 eligibles who filed a supplemental). Because of this, even a small number of accurate estimates may appear to represent a large frequency.

Independents had lower frequencies of discrepancy in all but the smallest categories of overestimating; however, they had higher rates of discrepancy in all but the largest category of underestimating. The differences in frequency rates in most categories were quite small. The major exception to this was the category of positive (underestimating) discrepancies larger than \$2,000. This range included 38.1 percent of the dependents but only 20.3 percent of the independents. Again, this difference probably reflects the differences in income levels between the two groups. Dependents have more large discrepancies because the general range of incomes for this group is much higher than for the independents.

EXHIBIT 4.9: DISCREPANCIES IN ADJUSTED GROSS INCOME FOR THE ELIGIBLE STUDY GROUP

AGI DISCREPANCY RANGE	BEOG-REPORTED DEPENDENCY STATUS		
	TOTAL	DEPENDENT	INDEPENDENT
TOTAL STUDY GROUP	7016 100.0	6189 87.8	857 12.2
TOTAL POSITIVE DISCREPANCIES GREATER THAN +200	4933 70.3	4386 62.5 72.2	547 7.8 63.8
+2000 OR MORE	2521 35.3	2347 33.3 38.1	174 2.5 20.3
+1001 - +2000	1138 16.2	960 13.4 15.9	158 2.2 18.1
+501 - +1000	746 10.6	633 9.0 10.2	113 1.6 13.2
+201 - +500	531 7.6	426 6.1 6.9	105 1.5 12.3
+200	918 13.1	688 9.8 11.2	230 3.3 26.8
-201 - -500	168 2.4	144 2.3 2.3	24 0.3 2.8
-501 - -1000	242 3.4	227 3.2 3.7	15 0.2 1.6
-1001 - -2000	256 3.6	243 3.5 3.5	13 0.2 1.5
-2001 OR LESS	499 7.1	471 6.7 7.6	28 0.4 3.3
TOTAL NEGATIVE DISCREPANCIES LESS THAN -200	1165 16.6	1085 15.3 17.6	80 1.1 9.5

COUNT
TOTAL PERCENT
ROW PERCENT

BEST COPY AVAILABLE

Because 100 percent of the independent study group had incomes under \$7,000, there were actually only two categories of income level: under \$4,000 and \$4,000 to \$6,000. These two groups were quite different. The \$4,000 to \$6,999 group was quite small and had a large rate of non-match. Exhibit 4.10 displays the results of this analysis. The lower income group had higher rates of accuracy and underestimating than the second group, but a significantly lower rate of overestimating. Since overestimating causes the individual to lose a portion of the grant to which he/she is entitled, and since all of the discrepancies were in the -\$2,001 or less group, this discrepancy rate highlights the difficulty of accurate estimation that is inherent in any estimation procedures. If 19.4 percent overestimate, it is likely that a similar percentage of underestimators may be considered to also have had difficulty and failed to arrive at an accurate estimate.

While it was difficult to determine any relationship between income discrepancies and income level for independents because of the distribution of the population in the income categories, a clearer relationship existed for dependents (see Exhibit 4.11). Here, the accuracy rate decreased steadily with increases in income level, from 13.4 percent for the lowest category to 5.7 percent for the highest income group. Accuracy was achieved by a small percentage at each income level. Ninety percent of dependents had discrepancies between their estimate and their true income. Those in the highest income category had the highest rate of positive discrepancies (underestimating) and the lowest rate of overestimating. Those in the lowest income category also had a low rate of overestimating. For the categories of income between \$4,000 and \$12,499 fairly similar rates of accuracy and discrepancies of both types were found, although there was a slight trend for increased underreporting with each increment of income, and similar decreases in accuracy and overreporting. In these categories the overreporting

EXHIBIT 4.10: FREQUENCY OF ADJUSTED GROSS INCOME DISCREPANCIES
BY BEOG-REPORTED AGI - ELIGIBLE INDEPENDENTS

AGI DISCREPANCY RANGE	BEOG-REPORTED AGI					
	TOTAL	LESS THAN \$4,000	\$4,000- \$6,999	\$7,000- \$9,999	\$10,000+	
TOTAL STUDY GROUP	704 100.0	642 91.2	62 8.8	*	*	COUNT TOTAL % COLUMN %
TOTAL POSITIVE DISCREPANCIES	483 68.6	446 63.4 69.5	37 5.2 59.7	*	*	
+2001 OR MORE	170 24.1	133 18.9 20.7	37 5.2 59.7	*	*	
+1001 - +2000	131 18.6	131 18.6 20.4	*	*	*	
+501 - +1000	88 12.5	88 12.5 13.7	*	*	*	
+201 - +500	94 13.3	94 13.3 14.6	*	*	*	
+200 - -200	187 26.6	174 24.7 27.1	13 1.8 21.0	*	*	
-201 - -500	11 1.6	11 1.6 1.7	*	*	*	
-501 - -1000	*	*	*	*	*	
-1001 - -2000	11 1.6	11 1.6 1.7	*	*	*	
-2001 OR LESS	12 1.7	*	12 1.7 19.4	*	*	
TOTAL NEGATIVE DISCREPANCIES	34 4.8	22 3.1 3.4	12 1.7 19.4	*	*	

EXHIBIT 4.11: FREQUENCY OF ADJUSTED GROSS INCOME DISCREPANCIES
BY BEOG-REPORTED AGI - ELIGIBLE DEPENDENTS

AGI DISCREPANCY RANGE	BEOG-REPORTED AGI					
	TOTAL	LESS THAN \$4,000	\$4,000- \$6,999	\$7,000- \$9,999	\$10,000- \$12,000	\$12,500+
TOTAL STUDY GROUP	6007 100	1966 32.7	1579 26.3	1496 24.9	737 12.3	229 3.8
TOTAL POSITIVE DISCREPANCIES	4358 72.5	1506 25.1 76.6	1080 18.0 68.4	1041 17.3 69.6	533 8.9 72.3	198 3.3 86.5
+1001 OR MORE	2544 39.0	812 13.5 41.3	589 9.8 37.3	531 8.8 35.3	296 4.9 40.2	116 1.9 50.6
+1000 - +2000	980 16.3	340 5.7 17.3	230 3.8 14.6	229 3.8 15.2	131 2.2 17.8	52 0.9 22.7
+501 - +1000	633 10.5	179 3.0 9.1	141 2.3 8.9	199 3.3 13.3	84 1.4 11.4	30 0.5 13.1
+201 - +500	401 6.7	175 2.9 8.9	120 2.0 7.6	84 1.4 5.6	22 0.4 3.0	*
+200 - -200	640 10.7	264 4.4 13.4	153 2.6 9.7	148 2.5 9.9	62 1.0 8.4	13 0.2 5.7
-201 - -500	113 1.9	36 0.6 1.8	25 0.4 1.6	52 0.9 3.5	*	*
-501 - -1000	210 3.5	44 0.7 2.2	72 1.2 4.5	69 1.1 4.6	25 0.4 3.4	*
-1001 - -2000	223 3.7	64 1.1 3.3	63 1.0 4.0	61 1.0 4.1	35 0.6 4.7	*
-2001 OR LESS	463 7.7	52 0.9 2.7	186 3.1 11.8	125 2.1 8.3	82 0.6 4.7	18 0.3 7.9
TOTAL NEGATIVE DISCREPANCIES	1009 16.8	196 3.3 9.9	346 5.8 21.9	307 5.1 20.3	142 2.4 19.3	18 0.3 7.9

averaged 21 percent of the study group, which was similar to the rate found for the over \$4,000 category of independents. Differences in income level, however, did not have a completely linear relationship to errors in estimations. It appeared that dependency status did have an effect on accuracy, while the effect due to income level was considerably weaker.

Reasons for Filing Supplemental

In order to determine if the reason for filing a supplemental application is in anyway related to the accuracy of estimation, the AGI discrepancy range was compared to the four supplemental reason categories. The four categories are:

- Death
- Loss of employment
- Unable to pursue income activities
- Separation or divorce

Analysis of this variable was confounded by two factors: 1) the categories of death and separation/divorce had extremely small frequency (14.2 percent and 1.6 percent of the study group respectively), and 2) these two categories had extremely high non-match rates (79.9 percent and 73.5 percent, respectively). Thus, conclusions drawn from the statistics of these two groups must be made with extreme caution. Underestimation was 79.5 percent for the death category and 88.4 percent for the separation/divorce category, both discrepancy rates being higher than average. Exhibit 4.12 presents the results of this analysis.

The supplemental reason, loss of employment accounted for 52.8 percent of the study group. The accuracy rate for this category was 10.1 percent which was somewhat lower than for the whole group (12.5 percent). Discrepancies were slightly more frequent for this group, in both directions than was true for the whole study group. Large, positive discrepancies greater than \$2,000 accounted for

EXHIBIT 4.12: FREQUENCY OF ADJUSTED GROSS INCOME DISCREPANCIES BY SUPPLEMENTAL REASON - ELIGIBLE STUDY GROUP

AGI DISCREPANCY RANGE	SUPPLEMENTAL REASON				
	TOTAL	DEATH	LOSS OF EMPLOYMENT	UNABLE TO PURSUE INCOME ACTIVITIES	SEPARATION OR DIVORCE
TOTAL STUDY GROUP	6875 100.0	288 4.2	3631 52.8	2884 41.4	112 1.6
TOTAL POSITIVE DISCREPANCIES GREATER THAN +200	4887 71.1	229 3.3 79.5	2643 38.4 72.8	1916 27.4 67.4	99 1.4 98.4
+2001 OR MORE	2520 36.7	104 1.5 36.1	1466 21.3 40.4	904 13.2 31.8	46 0.7 41.1
+1001 - +2000	1126 16.4	64 0.9 22.2	353 5.1 15.2	493 7.2 17.5	16 0.2 14.3
+501 - +1000	746 10.8	43 0.6 14.9	363 5.3 10.0	303 4.4 10.7	37 0.5 13.0
+201 - +500	495 7.2	18 0.3 6.3	261 3.8 7.2	216 3.1 7.6	*
+200	856 12.5	21 0.3 7.3	366 5.3 10.1	456 6.6 16.0	13 0.2 11.6
-201 - -500	154 2.2	*	81 1.2 2.2	73 1.1 2.6	*
-501 - -1000	242 3.5	10 0.1 3.3	131 1.9 3.6	101 1.5 3.5	*
-1001 - -2000	253 3.7	13 0.2 4.5	140 2.0 3.9	100 1.5 3.5	*
-2001 OR LESS	483 7.0	15 0.2 5.2	170 2.5 7.4	198 2.9 7.0	*
TOTAL NEGATIVE DISCREPANCIES LESS THAN -200	1133 16.5	38 0.6	622 9.0	473 6.9	*

4.38

BEST COPY AVAILABLE

40.4 percent of this category, while the total for the study group in this discrepancy range was 36.7 percent.

Persons citing unable to pursue income activities were the most accurate group, with 16 percent of the group having accurate estimations. Except in the largest category of underestimation, rates of discrepancies in all categories were similar to the rates for the total study group. Only 31.8 percent of this group had underestimated by more than \$2,000, while more than 36 percent of all other groups had underestimated by more than \$2,000.

Overall, the variations in the frequency for each supplemental reason in discrepancy range were less than 10 percent from the group with the lowest frequency to that with the highest. The only group with exceptions to this trend was those citing separation or divorce as the supplemental reason. Because this category represented such a small portion of the total study group its variations were not considered significant.

Filing Status - applicants/parents reporting divorced/separated/
widowed and filing either a joint or separate return to IRS

This analysis was attempted in order to examine whether persons in the categories of divorced/separated/widowed were more accurate when filing a separate return than filing a joint return with their former spouse. It was hypothesized that individuals filing separately would have more accurate estimates because the separate return is based on the income of the person providing the applicant's support for the application year. If the applicant/parent files a joint return with the former spouse, the IRS master file will have only the joint income figure on the individual's record. Therefore, the estimate, which should be based only on the person's income which supports the applicant, will appear to be substantially more discrepant from the IRS data for joint filers.

The analysis results were not meaningful because too few applications/parents in the divorced/widowed/separated category filed a

joint return, and the non-match rate for this group was very high. Because the confidentiality algorithm was utilized, these findings could not be interpreted.

Supplemental Signature Date

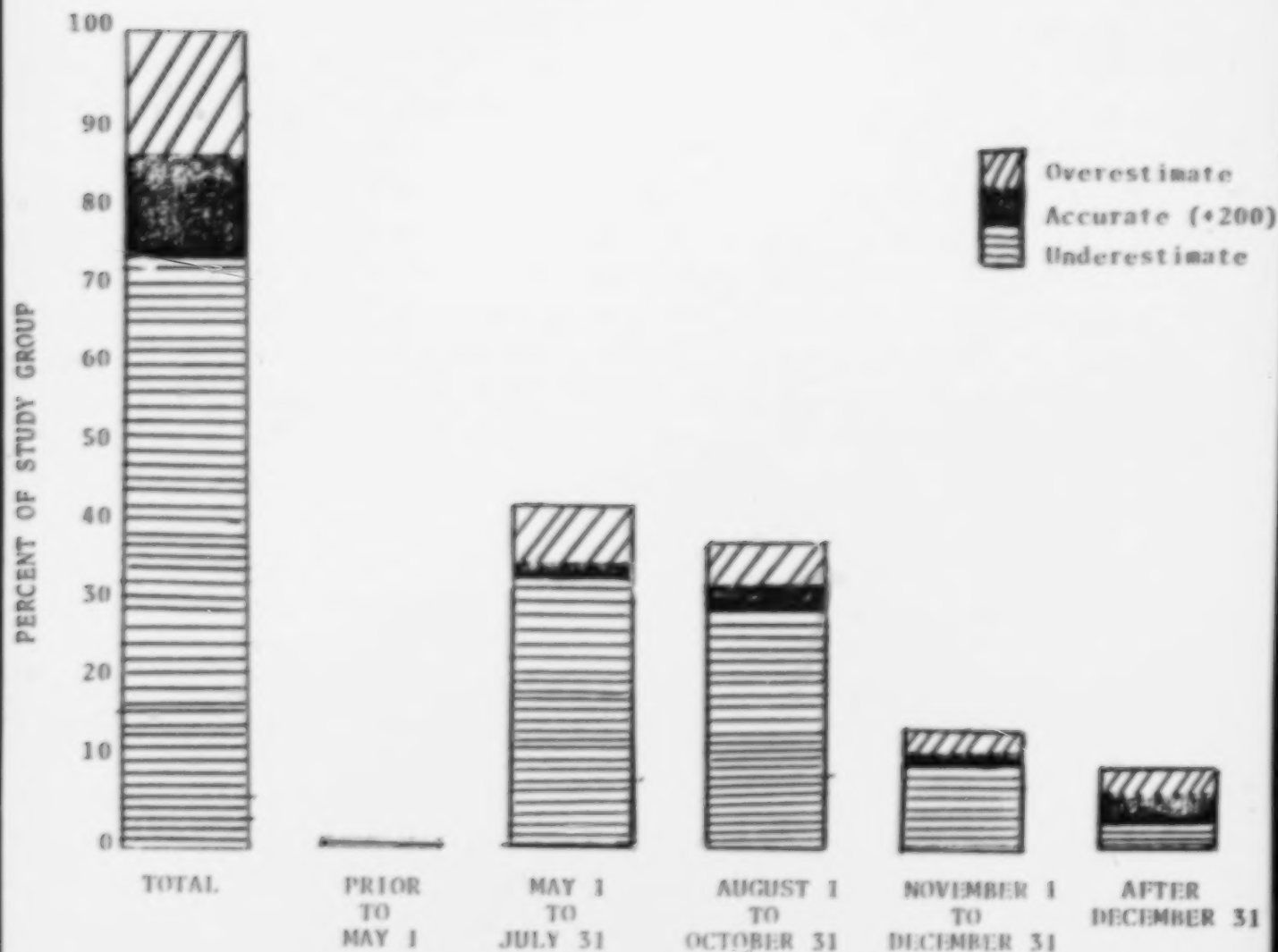
Supplementals, like regular applications, may be filed at any time during the application year. For Basic Grants, the estimation process involves adding the known income already received for the year to an estimate of the income expected for the remainder of the year. Thus, it was hypothesized that the date on which the estimate was made would have an effect on the accuracy of that estimation. (See Exhibit 4.13.)

For example, a person filing a supplemental in February would know his/her income for January of that year and part of February, but would estimate the income for the remainder of February and for March through December - or about 10 1/2 months. It can be expected that if another person were to file a supplemental in November, he/she would know the income for each month, January through October, and for part of November, and would have to estimate only for the remaining part of November and December - or about 1 1/2 months. Obviously, the latter person's estimate for the whole year is likelier to be accurate - with only 1 1/2 months to estimate, than the former person's estimate with 10 1/2 months income estimated.

The present analysis investigated the extent to which the date of filing the supplemental affected the accuracy of estimates. Did accuracy increase as the year progressed? A cross-tabulation was performed on the adjusted gross income discrepancy range and supplemental filing date.

The data suggest that the date of filing is closely related to the accuracy rate. Few people filed supplementals prior to May 1. After May 1 estimating accuracy increased steadily. For the period May 1 to July 31, accuracy was only 6.4 percent of the study group filing on that date. For those filing or submitting

EXHIBIT 4.13: FREQUENCY OF DISCREPANCIES IN ADJUSTED GROSS INCOME BY THE DATE OF SUPPLEMENTAL SIGNATURES OF APPLICANTS - ELIGIBLES



a correction after December 31, accuracy accounted for 48.5 percent of known cases. It is important to note, with this latter observation, that those filing after December 31 should have had a 100 percent accuracy rate, since the applicant would have had all income data for the year after the end of the year. The August 1 to October 31 period had slightly less overestimating than expected. Accuracy for this period was low as well.

Underestimation as measured for the total study group was over 70 percent for all date categories after May 1 except for the after December 31 category where the rate dropped to below 40 percent. Interestingly, as accuracy increased, and underestimating decreased, the non-match rate increased. Except for the prior to May 1 category, the match rate closely patterned the accuracy rate.

Question 2.1: Summary

According to the findings of this study thus far, it is clear that the date of filing has shown the strongest relationship to accuracy of estimation. Among other categories of applicants, the rate of underestimation was high for the population as a whole and for each specific characteristic, but no category had unusually skewed distributions of AGI discrepancies.^{1/} Date of filing was expected to show higher discrepancy rates because the applicant is reporting something that is not yet known to him/her. It was expected that later dates of filing supplementals would yield higher rates of accuracy. This was supported by the study results.

^{1/} Income level for dependents also showed an apparent relationship to estimating accuracy. This relationship, however, was not as definite as that found for filing date.

Question 2.2: Which category of applicants reflect higher or lower discrepancy rates in taxes paid?

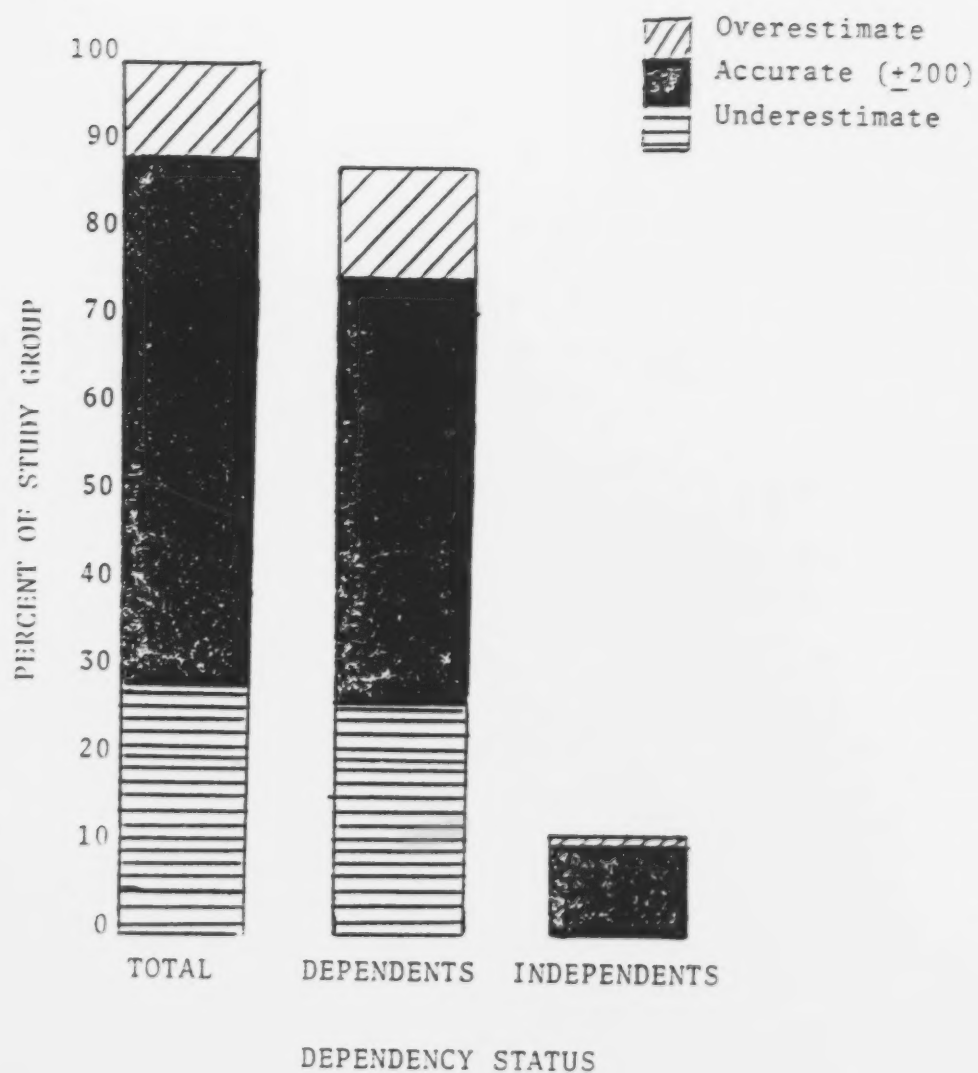
The taxes paid variable included in the Basic Grant data base for the estimated income population is calculated by the processing contractor from the (estimated) income provided on the supplemental application. The same formula is used by Basic Grants as is used by the Internal Revenue Service. Because Basic Grants does the calculating, it is assumed that no misreporting has occurred in the process of defining an individual's taxes paid. However, it is of interest to investigate the extent of inaccuracy in taxes paid, because it is included in the formula for determining BEOG eligibility index, along with AGI. Thus, the main query here is: what is the effect of inaccuracy in AGI on the accuracy of the taxes paid variable, and as a result, on the eligibility of applicants?

It should be noted that some degree of positive discrepancies (BEOG data lower in value than IRS data) was due to a definitional difference between the BEOG tax variable and the IRS tax variable. Basic Grants request that the applicant report only Federal income taxes paid. The IRS total acceptable tax includes Federal income taxes paid, as well as tax credits and taxes other than Federal income tax. Consequently, a small portion of the discrepancies may be attributed to those definitional variations.

In order to investigate the extent of discrepancies in taxes paid, a cross-tabulation analysis was performed between the taxes paid discrepancy range and dependency status. Exhibit 4.14 illustrates this analysis.

As may be seen, positive discrepancies accounted for 29.3 percent of the study group. Because the amount of taxes paid by a student or his/her family detract from their ability to pay for the student's education, it is to the applicants advantage to have a higher rate of taxes paid on the BEOG application than on IRS.

EXHIBIT 4.14: FREQUENCY OF DISCREPANCIES IN TAXES PAID BY
DEPENDENCY STATUS - ELIGIBLES



4.44

However, this 29.3 percent had a lower figure in the BEOG data base. Presumably the low tax figures in the BEOG data are the results of underestimation of income by the applicants. Since the estimates of income were lower than the actual income, the applicants were more able to pay for their education than they had anticipated when estimating. The fact that taxes paid had entered into the eligibility index and reduced the size of these students' awards slightly helped to limit the amount of Grant monies given to students who were discrepant in their income estimates.

An additional 58.6 percent of cases had accurate taxes paid (within ± 200). Negative discrepancies accounted for only 12.1 percent of the study group. Both of these results indicate that taxes paid figures in the Basic Grant data base were more accurate than the AGI estimates. This may be attributed to two factors. One, the range of taxes paid is considerably smaller than the range of income. Consequently, smaller discrepancies in taxes than in income should be expected. Two, the Basic Grant processor calculated taxes paid correctly for the marital status, income and number of exemptions reported.

In the comparison of the independent and dependent BEOG applicants it was found that the accuracy rate for independents was significantly higher (83.3 percent versus 55.2 percent) than for dependents. This may, in fact, be due to the lower income levels of independents which may result in lower taxes or none at all. The accuracy rate for taxes paid was higher for independents than was their accuracy rate for AGI.

Question 2.3: Which categories of applicants who estimate their incomes reflect higher or lower discrepancy rates in household size/exemptions?

Household size is the data item similar to IRS exemptions that is used in determining Basic Grant eligibility. Differences in the range of -1 to +1 between the IRS and BEOG data was considered

insignificant to account for the definitional incompatibility. Household size was examined in a cross-tabulation with dependency status. The household size reported to BEOG is the data element which the calculation of the eligibility index uses, unless the applicant has not reported this figure. In the latter event, exemptions are used in determining eligibility. Discrepancies in exemptions were cross-tabulated with number of exemptions reported to BEOG to determine if there is some relation between the number reported and discrepancies.

Accuracy in both analyses was found to be extremely high for the entire study group. (Non-matches were over 35 percent in both analyses). Independents were found to be somewhat more accurate in reporting household size than dependents, with a 96.1 percent incidence of insignificant discrepancies as compared to the dependents' 91.9 percent. All discrepancies, except for 0.3 percent, were negative for household size. That is, applicants overreported the number of persons in their household to Basic Grants, thus, potentially increasing the size of the award for which they were eligible.

The analysis of the relationship between exemption discrepancies and the number of exemptions reported to BEOGs showed a similar rate of accuracy, but those persons who were discrepant were primarily underreporters with only 0.4 percent of the group having overreported. It was found that those reporting either 1 or 2 exemptions to Basic Grants had the highest rates of underreporting (12.5 and 10.1 percent, respectively), lower rates of accuracy (less than 90 percent) and lower match rates (less than 50 percent). Interestingly, it would be expected that many independents would be among those reporting 2 or less exemptions, but independents as a whole were quite accurate in reporting household size.

Individuals who underreported exemptions to BEOG might have been eligible for larger awards if they had accurately specified the number of exemptions to which they were entitled.

Only the group reporting 3 exemptions had a measurable amount of overreporting, that is, reporting more exemptions to Basic Grants than to IRS. This amount was quite small however, only 2.3 percent of all persons reporting 3 exemptions to BEOG and 0.4 percent of the total study group.

Thus, only 0.4 percent of the study group would have become eligible to receive a significantly larger Grant than they were entitled to because they misreported their exemptions to BEOG. However, the exemption variable is only used by BEOG in the eligibility index when the household size is not reported.

To summarize, the household size/exemption variables showed a very high rate of accurate reporting, with dependents and those reporting 1 or 2 exemptions having the lowest rates of accuracy.

Question #2: Summary

From the analysis of discrepancies in three data elements used in the calculations of applicants' eligibility for Basic Grants, it was seen that AGI discrepancies are the most prevalent, with more than 70 percent of those studied having positive discrepancies (underestimates) and almost 20 percent having negative discrepancies (overestimates). Only 13 percent were able to accurately estimate their income. By contrast, 58 percent of the eligible applicants' taxes paid were accurate: the bulk of taxes paid discrepancies were such that they reduced the size of the applicants' potential Grant. Household size was reported accurately (within ± 1) 92 percent of the time. The household size discrepancies were positive, indicating that applicants reported more to Basic Grants than they had to IRS, and thereby potentially increasing the size of their Grant.

In general, independents were found to be more accurate than dependents and persons with low incomes tended to be more accurate than persons with higher incomes. Persons citing unable to pursue income activities tended to be the most accurate. However, the difference among groups were rarely as much as 10 percentage points different, and the non-matched population, not included in the analysis, was large enough to affect the distributions of estimating discrepancies. For this analysis it was assumed that the non-match was distributed randomly. The applicant characteristic which affected the accuracy in the most significant manner was the date of filing. Accuracy increased significantly as the last date of filing became later in the year. However, although accuracy for those filing after December 31 was greatly increased over the earlier time periods studied, there was still a significant number of discrepancies. Since the applicant should have had knowledge of the total year's income at this point, any supplemental or application correction should have reflected this knowledge. Apparently, the applicants, who were discrepant and who submitted applications in this time period, were correcting something other than income and did not also correct the income amount when submitting a revised student eligibility report.

QUESTION #3: DOES THE FREQUENCY OF CORRECTIONS MADE BY APPLICANTS TO THEIR BEOG APPLICATIONS RELATE IN ANY WAY TO MISREPORTING?

The Basic Grant application process allows applicants to change information reported if a mistake has been made. Corrections are entered into the BEOG master file as "trailer" to the applicants' records. The eligibility index contained in the latest trailer is the basis for determining the amount of the award. Because no supporting documentation has been requested by Basic Grants to verify corrections at the time they have been made, there has been no method of determining whether these corrections are justifiable. If the correction process is without verification the potential exists for

both intentional and unintentional use of this process as a means of misreporting. Applicants who file supplemental applications in order to estimate their incomes often file supplementals after having filed a regular application. Thus, the majority of estimators will have more than one trailer on their BEOG records. Again, no verifying information is required. The question which arises is whether the correction process is used by the applicant to increase the size of his/her award. Two aspects of the correction process were investigated in order to provide insight as to the use of application corrections:

- 3.1) Does the frequency of corrections relate to the size of AGI discrepancies?
- 3.2) Do any applicants report accurate AGIs, then submit corrections for supplementals in order to increase the size of their awards?

The findings of the analysis of these two questions is discussed below. A summary of findings related to Question 3 follows this discussion.

Question 3.1: Does the frequency of corrections relate to the size of AGI discrepancies?

In order to investigate this question, analyses were performed based on both AGI estimates provided to BEOG and the reasons applicants gave for filing the supplemental applications. Cross-tabulations were performed with these variables and the AGI discrepancy between estimates - IRS records of "true" income.

In the estimated income study group, 87.8 percent were found to have only one or two trailers, indicating that they had probably not filed corrections after submitting the supplemental. The majority of estimated income applicants filed their supplemental application as a correction to the original. Of these, approximately 70 percent in most categories underestimated their incomes on the 1974 application.

Accuracy was slightly lower for the group with one trailer than for the group with two trailers (10.9 percent and 11.6 percent, respectively). Apparently, this second trailer, then, does represent the filing of the supplemental as a correction to the original application. However, the positive discrepancies for the two-trailer group (72.5 percent) were slightly higher than for the one-trailer group. Increases occurred in the death, separation/divorce, and unable to pursue income activities categories. Negative discrepancies decreased from 18.7 percent for the one-trailer group to 15.9 percent for the two trailer group. Thus, there was some tendency towards estimates more "favorable" to the applicant for those filing two trailers, including a shift from negative (overestimating) discrepancies to accurate estimates.

For those filing three or more trailers, comprising 12.2 percent of the study group, there were no individuals occurring in the death and separation/divorce categories. Positive discrepancies for the other two categories increased substantially from an average of approximately 72 percent to 82.5 percent for three trailers and 80.9 percent for more than three. Negative discrepancies were reduced correspondingly, and the rate of accurate filers remained the same. Those reporting loss of employment had higher rates of discrepancies than those unable to pursue income activities. There was no consistent pattern for the separation/divorce and death categories because the sample sizes were too small for analysis.

There was no consistent relationship between income levels and the number of trailers on an eligible applicant's record. Those with one or two trailers had the same rates of underestimating, and those with accurate estimates increased from 9.1 to 12 percent. Negative (overestimates) discrepancies decreased from 17.4 to 14.5 percent in the income analysis. There was an increase in accuracy for the \$10,000-\$12,500 income category from 0 to 6 percent on the

second trailer, and similar increases for those in the under \$7,000 income categories, with corresponding decreases in both types of discrepancies. Underestimating increased by 13 percent for persons with three trailers, accuracy decreased slightly overall and several categories had no accurate reporting at all. Overestimating decreased for three or more trailers to 1.6 percent of the total. Only the \$10,000-\$12,500 category had any overestimating. The group with more than three trailers was quite small, about 102 individuals and analyses were inconclusive.

The non-match rates for one- and two-trailer populations remained the same for supplemental reason categories from one-to-two-trailer groups. It might also be noted that the size of the one-trailer group was smaller than that of the two-trailer group. This may be explained by the fact that many estimators file a supplemental after having filed a regular application.

Some of the inconclusive results obtained from this analysis were due to small numbers among the groups with the most trailers. This lack would indicate that these multiple corrections beyond the second were not used often by persons estimating their income in 1974. Among those who did utilize multiple corrections, the percentage of individuals with low incomes increased steadily and the percentage of those in higher income categories decreased similarly.

Question 3.2: Do any applicants report accurate AGIs, then submit corrections in order to increase the size of their awards?

Data on tables investigating the question of how many and what types of applicants filed accurate AGIs at any time prior to the latest trailer and were discrepant on the latest trailer represent too few cases to allow any generalizations to be made.

Another analysis of persons who were found to be ineligible based on a correction filed immediately before the supplemental (and were currently eligible) also resulted in an extremely small number of observations.

Question 3: Summary

In summary, it would appear that the correction process is not used by the estimated income population as a way of increasing their awards. Rather it seems that this process is used rarely, and only by those who need to make corrections.

QUESTION #4: DO APPLICANTS SUPPLY TRUE COPIES OF TAX FORMS WHEN PARTICIPATING IN THE VALIDATION EFFORT?^{1/}

The issue of whether supporting documentation provided by applicants participating in the individual validation effort was investigated by comparing discrepancy rates for that study and the IRS studies. If the discrepancy between BEOG-reported estimates and applicant-provided IRS forms was considerably and consistently different from the discrepancy between the BEOG-reported estimates and IRS data in the IRS master file, then some misinformation was provided by some of the applicants. Further, it was assumed that any false documentation provided would represent AGI data that were smaller than the true income figure in the IRS files. Thus, the applicant-provided IRS forms would be expected to result in smaller discrepancies when compared to the IRS provided data. In order to study this question, data on discrepancy rates obtained in the Individual Case Follow-up component of this study were compared to the data obtained in the IRS component. Overall frequencies of discrepancies obtained in both studies provided the major input to analysis.

For the Individual Case Follow-up Study, the discrepancy rate for underestimating (positive discrepancies) was found to be approximately 61.6 percent (with \$300 considered to be insignificant).

^{1/} This analysis was performed only on the eligible population.

For the IRS Sample, 71.1 percent of the study group (with 36.5 percent unknown) was found to be discrepant within \$200. If the discrepancy rate for the IRS data were based on the same \$300 figure for insignificant differences, the positive discrepancy rate would reduce to approximately 68.4 percent.

The difference found, approximately seven percent, is not considered to be significant given several factors regarding the sample. For instance, there was a 36.5 percent non-match rate in the IRS sample. The missing cases could affect the obtained rates of discrepancy by several percentage points. Also, the sampling error for the individual follow-ups was not known, and could affect any difference found. There was also a non-response rate of approximately 16 percent in the Individual Follow-up Study which could affect the results slightly.

In addition, many individuals in the Individual Case Follow-up component had discrepancies between their IRS data (provided by the applicants to BEOG) and their 1975-76 application. If the applicants had supplied false documentation, it would be expected that there would be few, if any, discrepancies between BEOG and IRS data. Therefore, it is concluded that the seven percent difference is not significant. This would mean that persons contacted using the Individual Case Follow-up methodology tend to submit accurate copies of their IRS forms when requested to do so.

QUESTION #5: WHAT ARE THE EFFECTS OF MISREPORTING BY THE ESTIMATED INCOME POPULATION ON BEOG ELIGIBILITY AND THE SIZE OF AWARDS?

The purpose of this analysis was to investigate the effect discrepancies found in BEOG-reported data had on the eligibility indices of the applicants studied. In order to do this, IRS versions of the major dependent variables of the Estimated Income Validation Study were substituted in the eligibility formula for the BEOG

reported data. For those applicants discrepant in one or more of these variables, it was of interest to determine if their awards would have been increased (or decreased) because of the multiple discrepancies.

The student eligibility index (SEI) is based on complex calculations involving the interrelationships of the applicant's family financial characteristics. The size of the index determines the amount of the applicant's award (the lower the index, the higher the award). A grant may be as high as \$1400 per year or one-half of the student's cost of education, whichever is the smaller amount.

The variables selected for examination were:

- Adjusted Gross Income
- Taxes Paid
- Household Size

These variables were also utilized in other analyses performed in this study, and have roughly equivalent meanings in both the Basic Grant and IRS data bases. A series of cross-tabulations were run substituting the IRS value for these factors into the eligibility index. First, only AGI was examined; then both AGI and taxes paid were substituted; and finally, all three variables were substituted into the index formula. For the purposes of this analysis, IRS data were considered to be "true" values of each variable. The IRS and BEOG definitions of household size were not strictly comparable, and this was recognized in formulating conclusions. However, it was felt that the substitution of household size on an equal basis would provide an indication of the effects of household size discrepancies on the eligibility index.

Obviously, using data which were discrepant with the data being replaced resulted in discrepancies in the new index calculations. It was these discrepancies which served as the dependent variable in this analysis. An eligibility index discrepancy was defined as:

The difference between the index as calculated using IRS data and that calculated using data reported to Basic Grants by the applicant.

Because the index resulting from IRS data elements was the criterion of accuracy, the index based on BEOG reported data was subtracted from the criterion index. The dependent variable was cross-tabulated with various applicant characteristics in order to attempt to determine the nature and extent of effect that index misreporting has on the Basic Grant Program. Specific tables were generated for the eligibility index discrepancies cross-tabulated with:

- Dependency Status
- BEOG-reported AGI
- BEOG-reported Student Eligibility Index
- AGI discrepancies

An analysis was also performed to examine the characteristics of those individuals who were eligible based on BEOG-reported data, but ineligible based on the substitution of the IRS data. The results of the cross-tabulations are presented below for the eligible study group and in the section 4.3.2.3 of this chapter for the ineligible study group. The eligibles were defined as those applicants who were eligible to receive a Basic Grant award on the basis of the data that they themselves provided on the BEOG Student Eligibility Report. The results presented below are primarily from the taxes paid substitution. The AGI, taxes paid and household size substitution was not emphasized because of definitional differences between BEOG and IRS versions of the household size variable. (See section 4.2 for discussion of these differences.)

SEI Discrepancies

All tables were calculated assuming that a difference smaller than ± 200 SEI points was an insignificant value, although each point could represent some award dollar value. This was done so that the

extent of major discrepancies could more easily be identified. Differences of 200 points are not likely to result in substantially larger awards. However, larger discrepancies can result in the payment of awards to individuals who do not really qualify for them. Also, because the individuals studied here had estimated their income, an index discrepancy of ± 200 could have represented any small errors in estimation made in good faith. Further, discrepancies could result from the inequality of data items between the data bases, or other factors not completely known at this time.

Based on these considerations, the Office of Education felt that the ± 200 range of discrepancies would be an appropriate limit to the "insignificant discrepancy" category. All SEI analyses below reflect this decision. Changing the "no-discrepancy" category (differences of less than ± 1) to the "insignificant discrepancy" category, had the following effects on the distributions for the eligible Estimated Income study group.

SUBSTITUTION	NUMBER OF INSIGNIFICANT DISCREPANCIES (Smaller than ± 200)	NUMBER OF DISCREPANCIES (Smaller than ± 1)	
AGI	3563	1696	(49%)
AGI, Taxes Paid	3629	1596	(44%)
AGI, Taxes Paid, House- hold Size	2896	1389	(48%)

It can be seen that nearly half of the insignificant discrepancy group actually had differences of less than ± 1 . Thus, the decision to create the ± 200 range appears to be supported by the study findings.

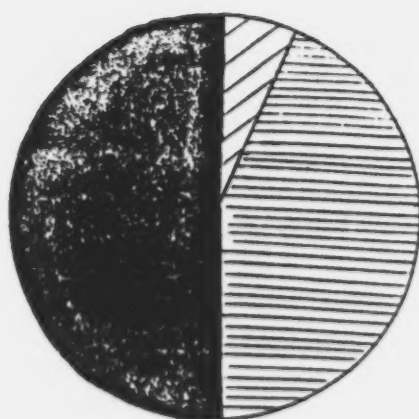
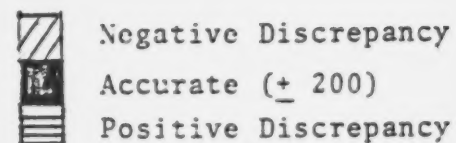
In general, the SEI discrepancies were smaller than those evidenced for the substitution variables themselves. For example,

negative discrepancies, or overestimation, were reduced from 16.6 percent found for adjusted gross income discrepancies, to approximately five percent for the SEI discrepancies. At the same time, the insignificant discrepancy rate of over 50 percent for each of the substitutions is substantially larger than the proportion of accurate estimators found for AGI (13.1 percent). Underestimators accounted for 70.3 percent of the discrepancies in the AGI analysis, but an average of 43.5 percent for the SEI discrepancy analysis. One result of the eligibility index analysis, then, is to show that for almost 30 percent of the eligible estimated income population found significantly discrepant in AGI estimates, the discrepancy obtained did not significantly affect their BEOG award. Exhibit 4.15 shows the distribution of the study group in each substitution analysis. As was true for other investigations, the non-match rate seriously impacted on study results. Non-match accounted for 35 percent of the population, including 31.4 percent of dependents and 53 percent for independents.

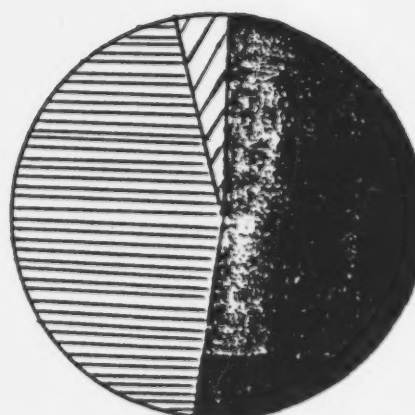
One important result of this analysis is that the largest numbers of the significant discrepancies appear in the largest categories for positive discrepancies (underestimators), but in the smallest discrepancy categories for negative discrepancies. That is, overestimators have smaller average discrepancies than underestimators. Applicants who underestimate, then, might be said to take "wilder guesses" when estimating for BEOG than do the overestimators.

Another characteristic of the underestimating group is that as the discrepancy categories get higher, the frequency drops substantially, except for the largest category of SEI discrepancies of 800 or more. The difference between the highest BEOG award and ineligibility is less than 1200 points. If a person has a discrepancy rate of more than 800 points, he/she will qualify for an award as much as two-thirds and more higher than that to which he/she is entitled. This category accounted for 18.9 percent of the AGI substitutions and 16.4 percent of the AGI and taxes paid substitutions.

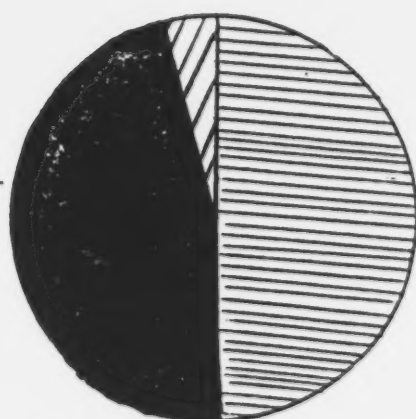
EXHIBIT 4.15: DISTRIBUTION OF DISCREPANCIES IN STUDENT
ELIGIBILITY INDEX FOR EACH SUBSTITUTION
ANALYSIS



AGI



AGI AND TAXES PAID



AGI, TAXES PAID,
HOUSEHOLD SIZE

Dependency Status

As noted in Question #2 above, dependents have a considerably higher match rate than independents, and represent a considerably larger portion of the study group, 88 percent to 12 percent. The eligible independents have a somewhat higher SEI accuracy rate based on both AGI and AGI and taxes paid substitutions. (See Exhibit 4.16.) The trend of accuracy favored the independents in both AGI and AGI plus taxes paid substitutions. Over half (57.3 percent) of the independent study group were accurate, while just 50 percent of the dependents were in this category in the AGI substitution analysis. For the AGI and taxes paid substitution, both groups increased slightly in accuracy. In both analyses independents had a smaller proportion of individuals with positive discrepancies.

Adjusted Gross Income

The adjusted gross income analysis was divided between dependents and independents since the income categories for each differed from one another. As was previously discussed, independents were slightly more accurate overall than were dependents. The purpose of the AGI analysis was to determine if the accuracy of the student eligibility index calculations is in anyway related to the income of applicants.

For dependents 82.5 percent of the population estimated that they had less than \$10,000 income for 1974. This large group was somewhat more accurate overall than those with over \$10,000 income. The low income categories had considerably fewer negative discrepancies than higher income groups. The lowest income category, those estimating under \$4,000 income for 1974, had the highest accuracy rates and lowest discrepancy rates of any group for both AGI and AGI plus taxes paid substitutions. However, 43.5 percent of this group were not included in the analysis; thus, there might be a considerable amount of error in the results. See Exhibit 4.17.

EXHIBIT 4.16: DISTRIBUTION OF STUDENT ELIGIBILITY INDEX DISCREPANCIES
BY DEPENDENCY STATUS AND TYPE OF IRS-SUBSTITUTION-
ELIGIBLE STUDY GROUP

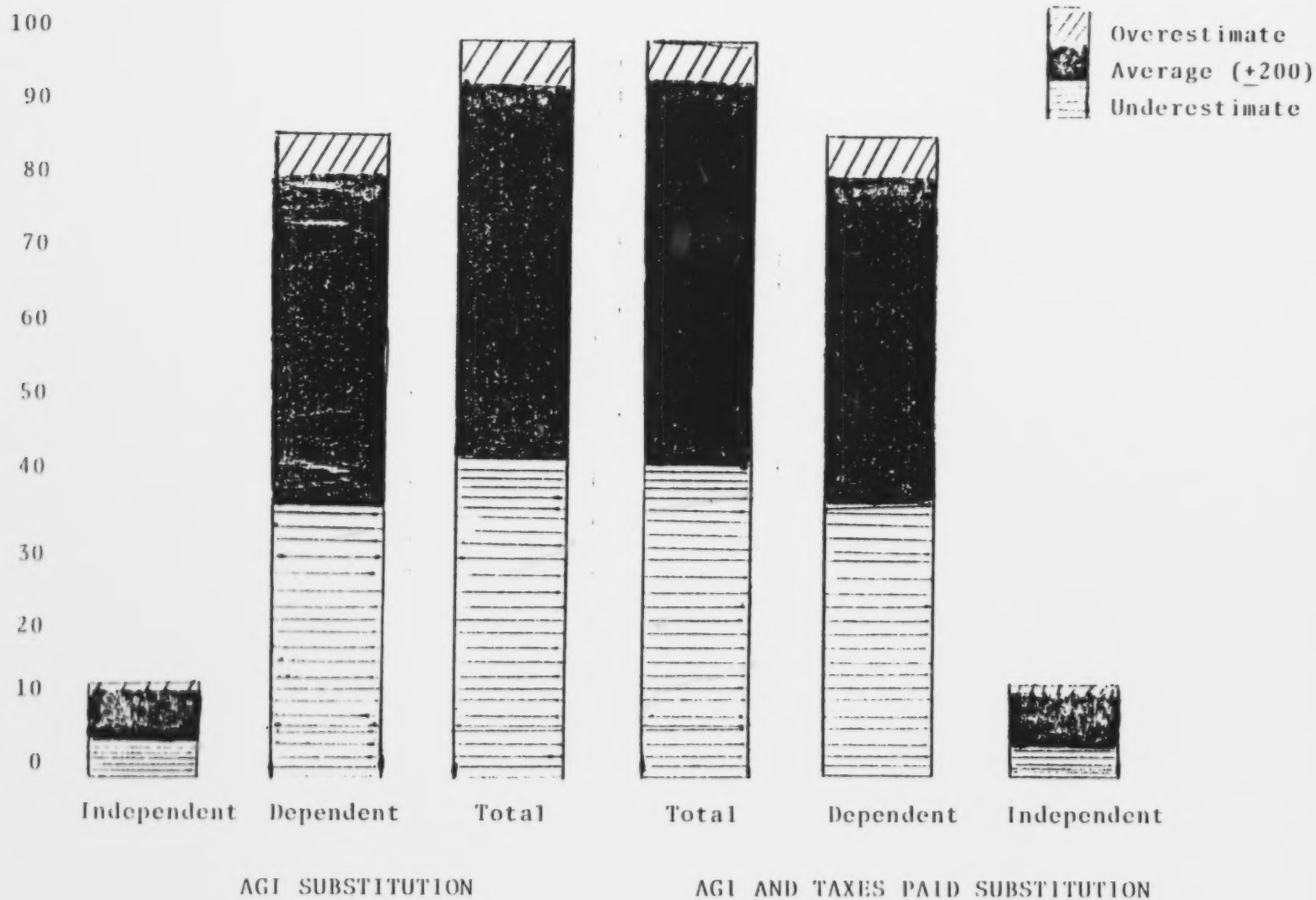


EXHIBIT 4.17: DISTRIBUTION OF STUDENT ELIGIBILITY INDEX DISCREPANCIES BY
LEVEL OF ADJUSTED GROSS INCOME - ELIGIBLE DEPENDENTS

SEI DISCREPANCY ^{1/}	BEOG-REPORTED AGI					
	TOTAL	UNDER \$4,000	\$4,000- \$6,999	\$7,000- \$9,999	\$10,000- \$12,4999	\$12,500 AND OVER
Total Study Group	6123	1972 32.2	1589 25.9	1490 24.3	764 12.5	308 5.0
Total Positive Discrepancies Greater than +200	2730 44.6	745 12.2 37.7	716 11.7 45.0	691 11.3 46.4	414 6.8 54.2	164 2.7 53.2
Total Accurate (+200)	3066 50.1	1215 19.8 61.6	811 13.2 51.0	669 10.4 44.9	257 4.5 33.6	114 1.9 37.0
Total Negative Discrep- ancies Less than -200	327 5.3	12 0.2 0.6	62 1.0 3.3	130 2.1 8.7	93 11.5 12.2	30 0.5 9.7

Count
Total %
Column %

^{1/}Based on Substitution of AGI

For eligible independents, too few individuals were found in the two income groups estimating their income to be \$7,000 or more, for analysis. There was virtually no overestimating evident for the independents. The non-match rate for independents overall also confounded income analysis since the small numbers of persons in this group resulted in a considerable amount of data suppression. Over 74 percent of eligible independents were in the under \$4,000 income category. On the whole, the under \$4,000 income category had higher accuracy and lower discrepancy rates than the over \$4,000 category. However, very few individuals were in the over \$4,000 category compared to those under \$4,000, and the higher income group had a non-match rate of 65 percent. In sum, it would appear that income as estimated by the applicant is not a variable which affects the accuracy rates of the independent applicants. These results are illustrated in Exhibit 4.18.

Student Eligibility Index

In order to determine the effect on Grant awards of discrepancies found through substitution of IRS data it was of interest to compare discrepancy rates for the levels of SEI obtained by the eligible estimating population through the information they reported to Basic Grants. Exhibits 4.19 and 4.20 show the cross-tabulations performed for AGI and AGI plus taxes paid substitutions.

Two related trends are evident in both tables. These are:

- As the SEI increased the rate of discrepancies increased;
- As the SEI increased the rate of accuracy decreased.

Over 53 percent of the two largest SEI groups were underestimators. At the same time the accuracy rate for the lowest SEI group was 64.1 percent for the AGI and 65.6 percent for the AGI and taxes paid substitutions. The lowest SEI group is also the largest, accounting for 42.7 percent of the total, while other categories accounted for an average of slightly over 11 percent each.

EXHIBIT 4.18: DISTRIBUTION OF STUDENT ELIGIBILITY INDEX DISCREPANCIES BY LEVEL OF ADJUSTED GROSS INCOME - ELIGIBLE INDEPENDENTS

SEI DISCREPANCY ^{1/}	BEOG-REPORTED AGI				
	TOTAL	UNDER \$4,000	\$4000- \$6,999	\$7,000- \$10,000	\$10,000 AND OVER
Total Study Group	797	671 84.2	126 15.8	*	*
Total Positive Discrepancies Greater than +200	333 418	267 33.5 39.8	66 8.3 52.9	*	*
Total Accurate (+ 200)	464 58.2	404 50.7 60.2	60 7.5 47.6	*	*
Total Negative Discrepancies Less than -200	*	*	*	*	*

Count
Total Percent
Column Percent

4.65

EXHIBIT 4.19: FREQUENCY OF DISCREPANCIES IN STUDENT ELIGIBILITY INDICES FOUND
IN AGI SUBSTITUTION ANALYSIS BY BEOG-REPORTED ELIGIBILITY
INDEX-ELIGIBLES

SEI DISCREPANCY	BEOG-REPORTED SEI						
	TOTAL	ZERO TO 200	201 TO 400	401 TO 600	601 TO 800	801 TO 1000	1000 TO 1200
Total Study Group	7035	3001 42.6	814 11.6	881 12.5	838 11.9	740 10.5	757 10.8
Total Positive	3085 43.9	1078 15.3 35.9	384 5.5 47.2	427 6.1 48.5	391 5.6 46.7	394 5.6 53.2	407 5.8 53.8
Total Accurate (+ 200)	3561 50.6	1923 27.3 64.1	374 5.3 45.9	354 5.0 40.2	366 5.2 43.7	258 3.7 34.9	286 4.1 37.8
Total Negative Discrepancies Less than -200	389 5.5	*	56 0.8 6.9	100 1.4 17.5	81 1.5 9.7	88 1.2 11.2	64 0.9 8.5

Count
Total %
Column %

4.64

EXHIBIT 4.20: FREQUENCY OF DISCREPANCIES IN STUDENT ELIGIBILITY INDICES FOUND
IN THE AGI AND TAXES PAID SUBSTITUTION ANALYSIS BY BEOG-REPORTED
SEI-ELIGIBLES

SEI DISCREPANCY	BEOG-REPORTED SEI						
	TOTAL	ZERO TO 200	201 TO 400	401 TO 600	601 TO 800	801 TO 1000	1000 TO 1200
Total Study Group	7026	3001 42.7	814 11.6	881 12.5	838 11.9	735 10.4	757 10.8
Total Positive Discrepancies Greater than +200	3054 43.5	1031 14.7 34.4	397 5.6 48.8	410 5.8 46.5	392 5.6 46.8	409 5.8 55.6	415 5.9 54.8
Total Accurate (+ 200)	3625 51.6	1970 28.0 65.6	367 5.2 45.1	378 5.4 42.9	367 5.2 43.8	260 3.7 25.4	283 4.0 37.4
Total Negative Discrepancies Less than -200	347 4.9	*	50 0.7 6.1	93 1.3 10.6	79 1.1 9.4	66 0.9 9.0	59 0.8 7.8

Count
Total %
Column %

4.65

The increase in the percentage of positive discrepancies in each SEI group was steady -- each larger than the last. Accuracy, however, was lowest for the 801-1000 point category and slightly higher over 1000.

The largest frequency of discrepancies occurred in the over 800 category for all discrepant applicants. Almost all of these would probably not have been eligible had their IRS data been reported instead of their estimates. This would have been especially true for those applicants whose estimate-based SEIs were greater than 401. For the highest SEI group virtually all positive discrepancies and some of the insignificant discrepancies would have resulted in ineligibility. However, these individuals qualified for minimal BEOG awards to begin with, and their discrepancies had less impact on the eventual disbursement of awards. Of course, the applicants discussed here have had discrepancies based on their estimates of income; they were expected to be considerably less accurate than the overall Basic Grant population.

Another important finding of this analysis found that only 4.9 percent of the study group with AGI substitutions were penalized by overestimates of income. The group affected most by overestimation was the 401-600 point category of SEI, with 10.6 percent. There were few if any, applicants with significant negative discrepancies for the 0-200 point SEI group.

Because it was so large, the 200-point SEI group bears close observation. First, although it has the lowest discrepancy rate of any group, those in this group who had positive discrepancies accounted for 14.7 percent of the study group in AGI substitution and 15.3 percent for the AGI plus taxes paid substitution study group. Secondly, the over 900-point discrepancy category of this group accounted for 6.3 percent and 7.4 percent of the AGI and AGI plus taxes paid substitution groups, respectively. Indeed, all categories of positive discrepancy and insignificant discrepancy for the under

200-point SEI group accounted for larger portions of the population than did other SEI groups. The discrepancy rate for this group was low, but the group itself was so large, that the frequency of discrepancies was actually the highest.

IRS - Ineligible Applicants

As was seen above, many individuals had discrepancies high enough to permit them to qualify for a Grant on the basis of their estimates although they were ineligible according to their actual incomes. An analysis of these individuals by dependency status and BEOG reported AGI shows that:

- 24.9 percent of the eligible estimated income study group would have been ineligible for a Grant based on their actual income for the year.
- 88 percent of this group were dependents, and 11 percent were independent.
- Over 75 percent of the independents in this group reported income of under \$4,000.
- Dependents in this analysis were almost evenly divided among income categories, but in the lower income categories comparatively fewer became ineligible than other categories of dependents, or than independents.
- The largest group who were ineligible based on IRS substitution were dependents in the \$7,000 to \$9,999 income category for which only AGI was substituted.
- AGI plus taxes paid substitution resulted in fewer ineligibles than either of the other substitutions. AGI substitution alone resulted in the most ineligibles, but only four more than AGI, taxes paid and household size.

In reviewing these results, it should be noted that independents are required by Basic Grant regulations to have substantially lower incomes to become eligible than are dependents. This accounts for the large percentages of eligible independents found in the lowest income categories. Thus, it would be expected that independents in low income categories (based on their own estimates) would account for the largest numbers of those considered ineligible after substitution of IRS data.

Another interesting result is that the proportions of dependents to independents in the groups found to be ineligible after data substitution was similar to that of the entire study group. Twelve percent of the ineligibles were from the independent category, and dependents accounted for 88 percent of the substitution-based ineligibles.

A possible explanation for the fact that AGI plus taxes paid actually resulted in the fewest ineligibles of the substitution analyses is that taxes paid had lower discrepancy rates than AGI. For estimated income applicants, taxes paid are calculated by Basic Grants, not by the applicant. Thus, the combination of AGI and taxes paid in substitution in the SEI resulted in fewer discrepancies than for AGI alone. The addition of household size raised the number of substitution-base ineligibles because each household size discrepancy had a significant impact on the calculation of an individual's eligibility. Exhibit 4.21 illustrates the eligibility analysis.

AGI Discrepancy Range

The final analyses performed for the eligibility index substitution analysis were a series of cross-tabulations with the purpose of determining if the AGI discrepancy ranges obtained for applicants were in any way related to the SEI discrepancy ranges. Because the objective was to delineate those discrepancies which were related to one another, small ranges of AGI discrepancies were utilized. This had the unfortunate effect of preventing detailed analysis because of the extent of the suppression of data in cells displaying less than ten individuals, since the eligible estimated income study group was comparatively small.

The largest numbers of cases were in the over \$2,000 AGI discrepancy ranges, particularly the positive discrepancies. Almost one-third (31 percent of the study group were in the 2000 or more category of AGI discrepancy. Of these, most were in the over \$00 point SEI discrepancy category. Approximately 16 percent of the

EXHIBIT 4.21: BEOG-REPORTED ELIGIBLE AND IRS-REPORTED INELIGIBLE
APPLICANTS WITH ESTIMATED INCOME BY BEOG-REPORTED
AGI AND DEPENDENCY STATUS

DEPENDENCY STATUS	BEOG-REPORTED AGI					
	TOTAL	LESS THAN \$4,000	\$4,000 TO \$6,999	\$7,000 TO \$9,999	\$10,000 TO \$12,499	\$12,500 OR MORE
TOTAL STUDY GROUP	1,628 100.0	484 29.9	342 21.1	370 22.8	291 18.0	133 8.2
DEPENDENT	1,429	331	304	370	291	133
Total Percent	88.2	68.5	88.9	100.0	100.0	100.0
Column Percent		20.4	18.8	22.8	18.0	8.2
INDEPENDENT	191	153	38	*	*	
Total Percent	11.8	31.5	11.1			
Column Percent		9.4	2.3			

NOTE: INDEPENDENT APPLICANTS WITH BEOG-REPORTED AGI EXCEEDING \$10,000
ARE REPORTED IN THE \$10,000 TO \$12,499 STRATUM.

ELIGIBILITY INDICES REPORTED IN THIS TABLE ARE BASED UPON SUBSTI-
TUTION OF THE IRS-REPORTED AGI, AND TAXES PAID FOR THE CORRESPOND-
ING BEOG-REPORTED DATA.

matched group was found in both the highest AGI discrepancy and the highest SEI discrepancy ranges. These individuals came almost entirely from the dependent applicant category, where 34 percent of the study group had positive SEI discrepancies with an AGI discrepancy range of +\$2,001 or more. Approximately half of these individuals, or 18 percent of the total dependents, had SEI discrepancies of +801 points or more. For the rest of the categories of both discrepancies, cases are more or less evenly distributed. Of the 50 percent insignificant SEI discrepancies, for example, no one size of AGI discrepancy is substantially different in frequency than any other. The data were not sufficient to permit further analysis, due to data suppression.

Question #5: Summary

To summarize, the major finding of the IRS data substitution, eligibility index analysis was that the student eligibility index had substantially fewer discrepancies than there were for adjusted gross income analysis. This was expected for two reasons. First, there are many factors involved in calculating SEIs, many of which lessen the direct input of income. For example, if both parents in a family work they will have a lower SEI, with all other factors being equal, than if the family income came from just one parent working. Secondly, many of the applicants had very small estimated incomes and small actual incomes for the application year. Therefore, even comparatively large discrepancies in AGI led to small discrepancies in SEI calculations.

Independents were found to be less discrepant in SEI than were dependents. This result was probably due to the effect of the relatively low incomes of eligible independents. Dependents were most discrepant in the high income, high AGI discrepancy ranges. Of all SEI discrepancy ranges, the over +800 point range had the most individuals except for the insignificant discrepancy range which accounted for approximately half of the cases. Other patterns that emerged for eligibles were that the frequency of eligibility index discrepancies

increased as size of the income estimate increased, and decreased as the size of the eligibility index increased.

4.3.2.3: Findings Obtained for the Ineligible Study Group

This section of the report discusses the accuracy of estimation for those Basic Grant applicants who were determined to be ineligible to receive a Grant based on the estimate of income they provided on their supplemental BEOG applications. The findings of analyses performed are presented in the order of the analysis questions listed in Section 4.1.

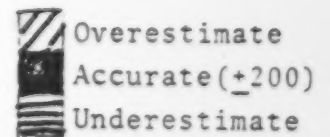
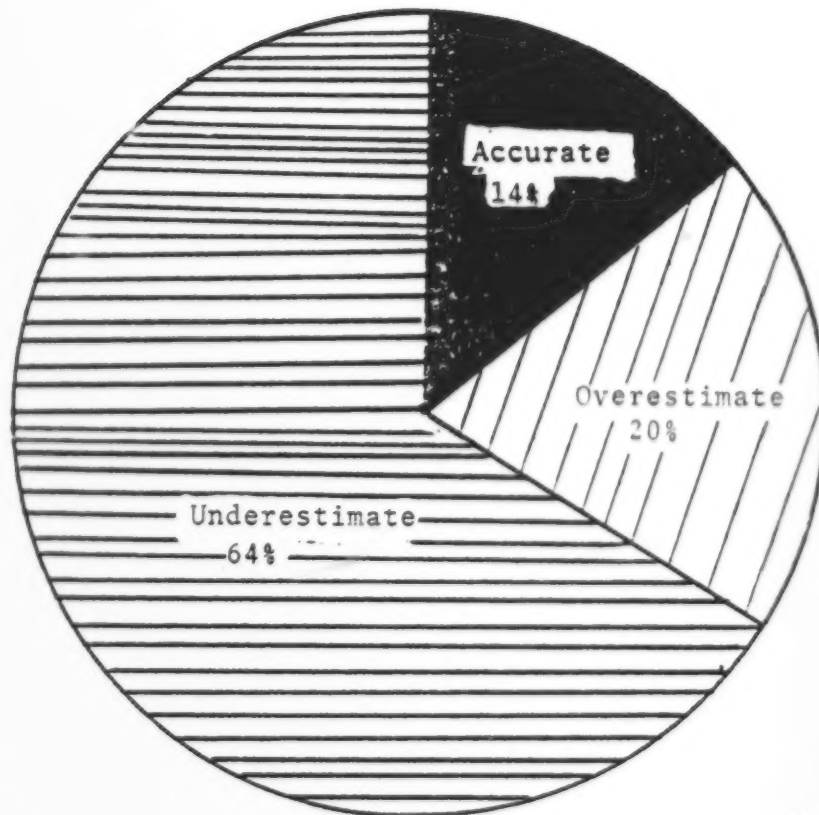
QUESTION #1: WHAT IS THE OVERALL ACCURACY OF ESTIMATION?

Of those persons who submitted supplemental applications to Basic Grants for the 1974-75 application year, about 38 percent were found to be ineligible. Just as poor estimating by eligibles may result in an inappropriate award, so too can poor estimating by ineligible applicants result in their eligibility. Any analysis of the errors made by eligible BEOG applicants, then, must also consider an assessment of the ineligibles.

Overall, 14 percent of the ineligible applicant study group were found to have estimated their income accurately. Discrepancies amounted to 84 percent of the total. Non-matches accounted for 33.8 percent of the ineligible population. As explained previously, it may be assumed, to some degree, that the unknown cases are distributed in the same proportion as the study group. The discrepancies were divided with 64 percent positive and 20 percent negative. Exhibit 4.22 illustrates the overall accuracy of the ineligible population.

When compared to the eligible estimated income population, it may be seen that the distribution of discrepancies among the two study groups was similar, with the ineligibles being slightly more accurate, having more overestimators, and having slightly fewer underestimators. There were also fewer non-matches than for the eligible population. Thus, it may be considered that, for the estimated income population, there are few and minor differences

EXHIBIT 4.22: OVERALL ACCURACY OF ADJUSTED GROSS INCOME
ESTIMATES - INELIGIBLES STUDY GROUP



in discrepancy rates between eligibles and ineligibles.

QUESTION #2: HOW IS ACCURACY OF ESTIMATION ASSOCIATED WITH
APPLICANT/PROGRAM CHARACTERISTICS?

Question 2.1: Which categories of applicants reflect higher or
lower discrepancy rates in adjusted gross income?

Dependency Status

As illustrated in Exhibit 4.23, it may be seen that, for the ineligible population, independents are substantially more accurate than dependents even when the high non-match rate for independents is considered. However, the independent population is only one-eighth the size of the eligible population.

The largest differences in the discrepancy rate between dependents and independents occur for the ± 2001 or more discrepancy ranges. For independents who estimated their family's income for 1974 the accurate estimates accounted for 32.3 percent of the study group; the remaining 56.2 percent were underestimators; and 11.5 percent were overestimators. For dependents, accurate estimators were 12.7 percent of the study group, with 66.3 percent underestimators and 20.9 percent overestimators.

In terms of income, the distributions of ineligible independents and dependents were considerably different from one another, as well as from the eligible population. As might be expected, there were higher frequencies of higher income persons in the ineligible population; they might have, if their incomes were lower, been eligible.

For independents, analysis of income characteristics was confounded because of data suppression which prevented the display of cells with less than 10 observations. Thus, with the small number of independents, much of the data were suppressed for the two income categories falling at \$7,000 and over. For lower incomes, there did not appear to be a difference between the

EXHIBIT 4.23: FREQUENCY OF ADJUSTED GROSS INCOME DISCREPANCIES
BY DEPENDENCY STATUS - INELIGIBLES

AGI DISCREPANCY RANGE	BEOG-REPORTED DEPENDENCY STATUS			
	TOTAL	DEPENDENT	INDEPENDENT	
Total Study Group	4327	3934	393	
Total Positive Discrepancies	2833	2612	221	COUNT
	65.3	60.3	5.1	TOTAL PERCENT
		66.3	56.2	COLUMN PERCENT
More Than -2001	1409	1340	69	
	32.6	1.0	1.6	
		34.1	17.6	
-1001 - -2000	691	646	45	
	15.0	14.9	1.0	
		16.4	11.5	
+501 - -1000	465	405	60	
	10.7	9.4	1.3	
		10.5	15.3	
-201 - +500	268	221	47	
	6.2	5.1	1.1	
		5.6	11.9	
+200	628	501	127	
	14.5	11.6	2.9	
		12.7	32.3	
-201 - 500	101	101	*	
	2.3	1.5		
		1.6		
-501 - -1000	136	136	*	
	3.1	3.1		
		3.5		
-1001 - -200	206	195	11	
	4.8	4.5	0.3	
		4.9	2.8	
Less Than -2000	423	389	34	
	9.8	9.9	0.8	
		9.9	8.7	
Total Negative Discrepancies	866	821	45	
	20.0	18.9	1.0	
		20.9	11.3	

4.74

BEST COPY AVAILABLE

under \$4,000 - \$6,999 categories. However, conclusions to the population were also limited for the low income categories since approximately 70 percent of these cases were also missing. (See Exhibit 4.24.)

Analysis of income categories for dependent applicants showed that the largest number have incomes over \$12,500, and this group had the lowest non-match rate. This may be compared to the eligible population where the opposite was true (the highest income had a high non-match rate). This group also had the highest rate of accuracy for ineligibles: 14.5 percent of the study group. Overall, the positive discrepancy rate was 69.5 percent. The largest percentage of discrepancies occurred in under \$4,000 category and was stable in all categories between \$4,000 and \$12,499. (See Exhibit 4.25.)

Of particular interest in the study of the ineligibles is the frequency of overestimators. Overall, this does not occur substantially more often than for the eligibles. However, in the highest income categories, the amount of overestimating is greater than is found elsewhere. This would indicate, particularly for those individuals who fall in the \$2,000 or less category of discrepancies, that they might have been eligible to receive a grant if they had been able to estimate more accurately. Overestimating is infrequent for the lowest income category.

Reasons for Filing Supplemental

In order to determine if the accuracy of estimation is related in any way to the reason given by applicants for filing a supplemental the AGI discrepancy range was compared to the four supplemental reason categories which are:

- Death
- Loss of employment
- Unable to pursue income activities
- Separation or divorce

4.75

EXHIBIT 4.24: FREQUENCY OF ADJUSTED GROSS INCOME DISCREPANCIES BY LEVEL OF AGI --
INDEPENDENT INELIGIBLES

AGI DISCREPANCY RANGE	TOTAL	BEOG-REPORTED AGI			
		LESS THAN \$4000	\$4,000 \$6,999	\$7,000+	
TOTAL STUDY GROUP	198	83 41.9	78 39.9	37 18.7	
TOTAL POSITIVE DISCREPANCIES GREATER THAN +200	105	54 27.3 65.1	51 25.8 65.9	* * *	COUNT TOTAL PERCENT COLUMN PERCENT
2001 OR MORE	37	11 5.6 13.3	26 13.1 33.3	* * *	
+1001 - +2000	13	* 6.6	13 6.6 15.7	* * *	
+501 - +1000	42	10 15.2 36.1	12 6.1 15.9	* * *	
+201 - +500	13	13 6.6 15.7	* * *	* * *	
+200	68	29 14.6 34.9	27 13.6 34.6	12 6.1 32.4	
-201 - 500	*	* * *	* * *	* * *	
-501 - -1000	*	* * *	* * *	* * *	
-1001 - -2000	*	* * *	* * *	* * *	
-2001 OR LESS	25	* 12.6	* * *	25 12.6 67.6	
TOTAL NEGATIVE DISCREPANCIES LESS THAN -200	25	* 12.6	* * *	25 12.6 67.6	

4.76

EXHIBIT 4.25: FREQUENCY OF DISCREPANCIES IN ADJUSTED GROSS INCOME BY LEVEL OF AGI --
DEPENDENT INELIGIBLES

AGI DISCREPANCY RANGE	TOTAL	LESS THAN \$4000	\$4000 \$6999	\$7000 \$9999	\$10,000 \$12,499	\$12,500+	COUNT TOTAL PERCENT COLUMN PERCENT
TOTAL STUDY GROUP	3712	561 15.1	564 15.2	687 18.5	604 16.3	1296 34.9	
TOTAL POSITIVE DISCREPANCIES GREATER THAN +200	2581 69.5	504 13.6 89.9	405 10.9 71.8	492 13.2 71.6	424 11.4 70.2	756 20.3 58.3	
+2001 OR MORE	1339 36.1	255 6.9 45.5	222 6.0 33.4	252 6.8 36.7	226 6.1 37.4	384 10.3 29.6	
+1001 - +2000	646 17.4	132 3.6 23.5	78 2.1 13.8	114 3.1 16.6	98 2.6 16.2	224 6.0 17.3	
+501 - +1000	405 10.9	72 1.3 12.8	90 2.4 16.0	75 2.0 10.9	52 1.4 8.6	116 3.1 9.0	
+201 - +500	191 5.1	45 1.2 8.0	15 0.4 2.7	51 1.4 7.4	48 1.3 7.9	32 0.9 2.4	
+200	382 10.3	42 1.1 7.5	42 1.1 7.4	66 1.8 9.6	44 1.2 7.3	188 5.1 14.5	
-201 - 500	62 1.7	*	*	12 0.3 1.7	22 5.9 3.6	28 0.8 2.2	
-500 - -1000	115 3.1	15 0.4 2.7	12 0.3 2.1	24 0.6 3.5	12 0.3 2.0	52 1.4 4.0	
-1001 - -2000	183 4.9	*	45 1.2 8.0	30 0.8 4.4	32 0.9 5.3	76 2.0 5.9	
-2001 OR LESS	389 10.5	*	60 1.6 10.6	63 1.7 9.2	70 1.9 11.6	196 5.3 15.1	
TOTAL NEGATIVE DISCREPANCIES LESS THAN -200	749 2.0	15 0.4 2.7	117 3.2 20.7	129 3.4 18.8	136 3.7 22.5	352 9.5 27.2	

Out of the total of 6,535 ineligible applicants who filed supplementals, a total of 89 individuals cited separation or divorce as their reason for estimating their income. Of these, 85 percent were non-matches, thus the analysis for this category of reason was not useful. For other categories some generalization can be made which illustrate the effect of the reason on the accuracy of estimation.

For death cited as the reason, 64 percent of the total sample was unknown due to non-matches. This category was a larger portion of the population than it had been for the eligibles. Discrepancies accounted for over 85 percent of those in this category, which was a higher rate than for the other categories. The death category also had a somewhat lower rate of accuracy than other categories (8 percent) which was also well below the overall rate of 13.9 percent for the total study group.

As was true for the eligibles, ineligible applicants citing either loss of income or unable to pursue income activities made up the bulk of the applicants. However, due to the larger rate of citations for death, both these categories are smaller proportions of the whole than was found for eligibles. Underestimating was somewhat less for ineligibles than eligibles in these two categories, while accuracy was slightly more frequent and overestimating somewhat more frequent for ineligibles. Those reporting unable to pursue income activities were, in general, more accurate than those citing loss of employment. These individuals also had a higher rate of overestimators, and a lower rate of underestimators than the loss of employment category.

In summary, it may be noted that the ineligible population tended to be more accurate than eligibles, but not by large differences. Those citing death as a reason for filing a supplemental had a very high non-response rate and a higher rate of discrepancies while accounting for 25.9 percent of all ineligibles. Overall, those citing unable to pursue income activities were

somewhat more accurate than those citing loss of employment but both groups were more accurate than the total study group. Ineligibles also had a higher rate of overestimators than did the eligibles, indicating that in some cases, they may have been eligible for a Basic Grant if their estimates had been more accurate. Exhibit 4.26 presents an illustration of these results.

Filing Status - applicants/parents reporting divorced/separated/widowed to BEOG and filing either a joint or separate return to IRS

The purpose of this analysis was to examine whether persons in these categories were less accurate when filing joint returns than when filing separately, and if so, to examine the discrepancy rates relative to AGI. It was hypothesized that individuals in the same situation, but filing separately would have more accurate estimates of their income, because the year's income figure would be based only on the one person's income, and the same income figure reported to BEOGs would be reported to IRS. For those individuals in the divorced/separated/widowed category who file joint return with their (former) spouses, the income reported to BEOGs would be that of the person providing the applicant's major source of support, while the figure reported to IRS would be both parent/applicant and (former) spouse. Thus the discrepancies would be quite large between the one income reported to BEOG and the two incomes in the IRS records. It was also expected that the non-match rate would be particularly high for the joint filers since matches could be made only on one parent's social security number, and the IRS records may have been stored under the other parent's number.

The analysis, in fact, showed, as was true for the eligibles, that too few persons who could have, did file a joint return. In fact, the total number of observations that were permissible was 24, the total population of persons in the divorced/widowed/separated category was 2,116 for ineligibles. Non-match rates even

EXHIBIT 4.26: FREQUENCY OF DISCREPANCIES IN ADJUSTED GROSS INCOME
BY SUPPLEMENTAL REASON - INELIGIBLES

AGI DISCREPANCY RANGE	SUPPLEMENTAL REASON				
	TOTAL	DEATH	LOSS OF EMPLOYMENT	UNABLE TO PURSUE INCOME ACTIVITIES	SEPARATION OR DIVORCE
Total Study Group	4280	602 14.1	2126 49.7	1539 35.9	13 0.3
Total Positive Discrepancies Greater than +200	2809 65.7	469 10.4 77.3	1405 32.8 66.1	922 21.5 59.3	13 0.3 100.0
+2001 or More	1408 32.9	213 4.4 35.3	743 17.4 34.3	439 10.3 28.5	13 0.3 100.0
+1001 - +2000	683 15.4	135 3.2 22.4	320 7.5 15.4	228 5.3 14.8	*
+501 - +1000	458 10.7	75 1.8 12.4	225 5.3 10.6	158 3.7 10.3	*
+201 - +500	260 6.1	46 1.1 7.6	117 2.7 5.5	97 2.3 6.3	*
+200	594 13.9	49 1.1 8.1	298 7.0 14.0	247 5.7 16.0	*
-201 - -500	89 2.1	14 0.3 2.3	48 1.1 2.3	27 0.6 1.8	*
-501 - -1000	150 3.5	24 0.6 3.9	49 1.1 2.3	77 1.8 5.0	*
-1001 - -200	214 5.0	21 0.5 3.5	112 2.6 5.3	81 1.9 5.3	*
-2001 or Less	424 9.3	25 0.6 4.2	214 5.0 10.0	188 4.4 12.2	*
Total Negative Discrepancies Less Than - 200	377 20.5	84 1.4 13.4	423 9.9 19.9	370 8.6 24.4	*

Count
Total Percent
Column Percent

among those filing separate returns were quite high; somewhat more than 75 percent of all applicants/parents in both categories were not included in the analysis.

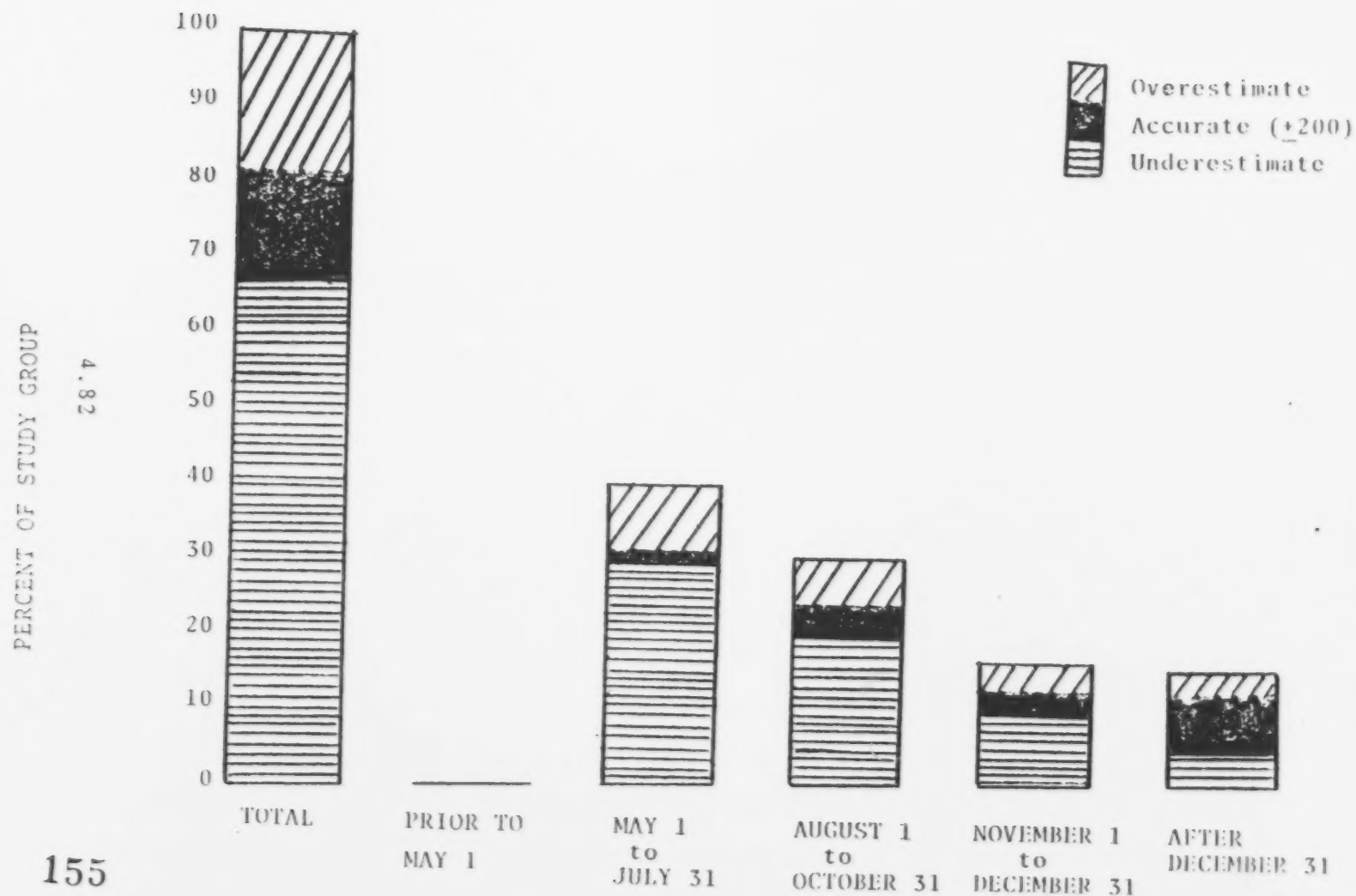
Supplemental Signature Date

Analysis of the accuracy of applicants as related to the date of the filing of the supplemental application reveals that the accuracy increased substantially for later filing dates with corresponding reductions in underestimates. Overestimation remained fairly consistent for all time periods except the last when it was reduced in size by a significant amount. Because of non-match and data suppression there were no cases in the period prior to May 1. The ineligible population exhibited similar characteristics of misreporting compared to the eligible in all categories.

Because the applicant who filed a supplemental recorded the amount earned to date and estimated the income for the remainder of the year to be affected by the change in circumstances, the result of increased accuracy for later filing dates was expected. The later in the year the application is filed the higher the proportion of known income (that already earned in the year) to unknown (that to be earned in the remainder of the year): this increases the chances of accuracy. Thus, persons who filed a supplemental after December 31 should have been 100 percent accurate in their estimate, since they had completed the year's earnings. That the individuals in this category remained substantially discrepant (48.6 percent of cases analyzed) indicated that there may have been some degree of intentional misreporting or errors not corrected on applications with signature dates after December 31. See Exhibit 4.27.

EXHIBIT 4.27:

FREQUENCY OF DISCREPANCIES IN ADJUSTED
GROSS INCOME BY DATE OF SUPPLEMENTAL
SIGNATURE



Question 2.1: Summary

For the ineligible Basic Grant application population, it appeared that the frequency of misreporting was strongly associated with the date of filing, as explained above, and there was also some relationship between accuracy and the reason given for filing a supplemental. For the date of filing, accuracy increased as the year progressed. The relationship of accuracy to the supplemental reason, was not as clear cut, but individuals citing death were declared ineligible for Grants more frequently than other categories, and the persons in this category underestimated more frequently than persons in other categories. There was also a small amount of greater accuracy among those citing unable to pursue income activities as compared to those citing loss of employment. The relationship between the supplemental reason categories and AGI discrepancy is stronger for the ineligible group than for the eligibles.

Question 2.2: Which category of applicants reflects a higher or lower discrepancy rate in taxes paid?

Discrepancy in taxes paid for the estimated income populations was cross-tabulated only with dependency status. Analysis was limited because taxes paid were not reported by the applicant, rather they were calculated using the estimate of income in the regular IRS tax formula by Basic Grants. Therefore, it was of interest to check the effect of AGI discrepancies on taxes paid, but more detailed analysis was not necessary. Some indication of discrepancy rates for taxes paid was useful since this variable entered into the eligibility index, and errors in AGI which resulted in errors in taxes paid could have compounded discrepancy ranges in the eligibility index. Question #5 discusses effects of misreporting on the eligibility index.

There was a fairly high degree of accuracy for taxes paid in the ineligible population. Examination of the study group indicates, that, overall, 46.4 percent of the cases had accurate

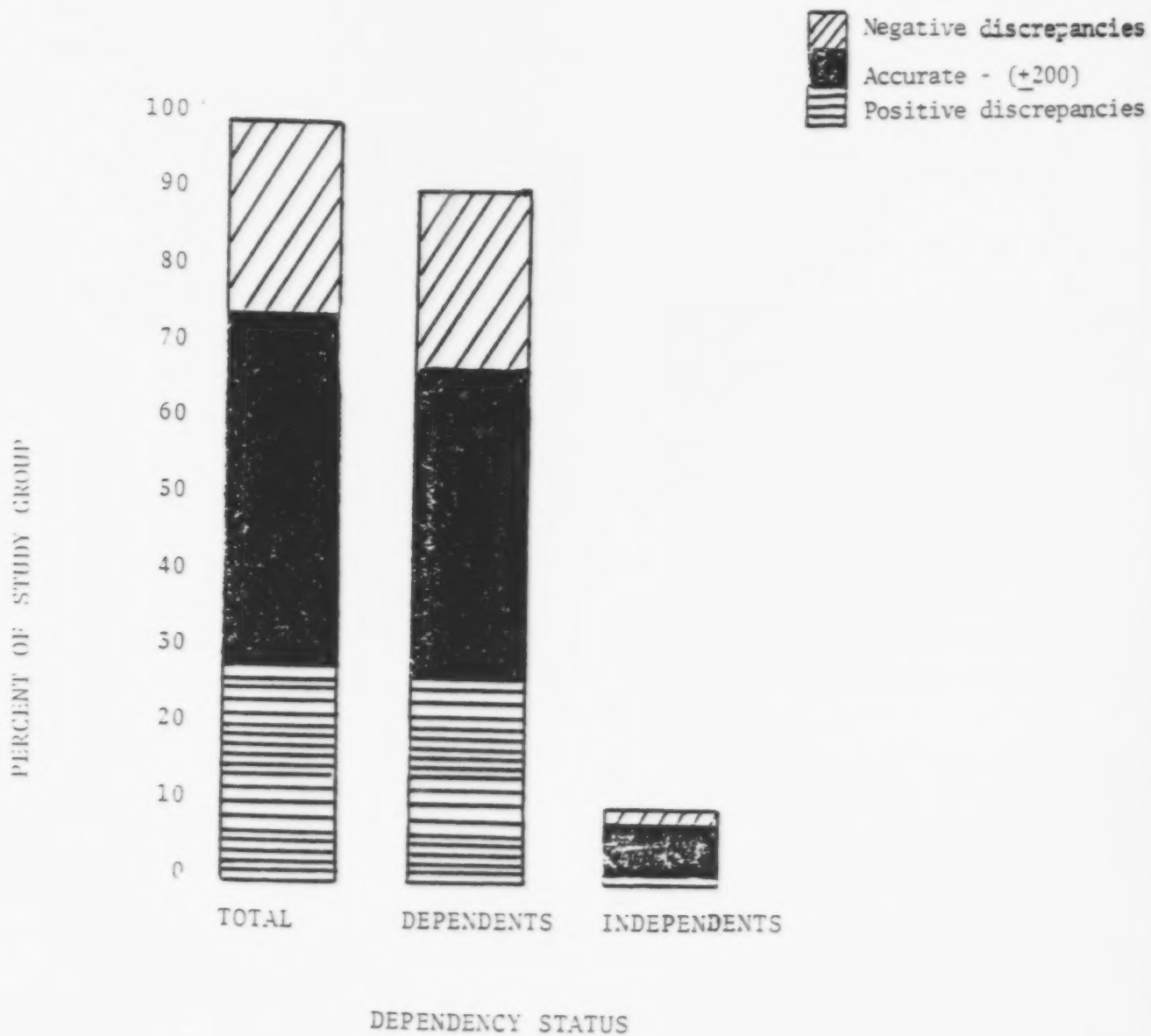
taxes paid. Significantly, the discrepancies found were almost evenly divided between positive and negative - at slightly over one-quarter of the study group each.

Thus, although the effect of errors in AGI may be compounded by taxes paid calculations utilizing the discrepant AGI, this effect was felt on only slightly more than one-half the applicants. This result probably occurred because of the low income categories of most of the applicants, hence small amounts of taxes paid. There was probably a large percentage of individuals in the population paying no taxes at all or a very small amount on their true income. An underestimate of AGI may then have resulted in a zero or small amount of taxes. Dependents were less accurate in taxes paid than independents and had higher rates of discrepancy. This result was most likely due to the higher income levels of dependents compared to independents; hence, larger amounts of taxes paid. Independents, however, had a much higher rate of non-matches. Exhibit 4.28 shows the results of the taxes paid analysis.

Question 2.3: Which categories of applicants who estimate their income reflect higher or lower discrepancy rates in household size/exemptions?

The analysis of household size and exemption discrepancies investigated the extent of the difference between the data reported to Basic Grants and that reported to IRS. Because of a definitional difference between the BEOG and IRS forms on the key variable, household size, which is included in the eligibility index calculation, discrepancy of ± 1 on either variable was considered insignificant. This being the case, discrepancies amounted to less than two percent of the study group for exemptions. For household size, discrepancies, usually overreporting, totaled less than 4 percent of known cases. Thus, it would appear that, in general, the estimated income population accurately reported their household size/exemptions to Basic Grants

EXHIBIT 4.28: FREQUENCY OF DISCREPANCIES IN
TAXES PAID BY DEPENDENCY STATUS-
INELIGIBLES



Even if the +1 range is not utilized and zero is the only non-discrepant category, the rate of accurate reporting was 80 percent or better for known cases.

QUESTION #3: DOES THE FREQUENCY OF CORRECTIONS MADE BY APPLICANTS TO THEIR BEOG APPLICATIONS RELATE IN ANY WAY TO MISREPORTING?

The BEOG application process includes a correction procedure to allow for errors made by applicants. The BEOG master file maintains records of all correction transactions and retains old records as well as changes. This is done by means of a variable-length record on the BEOG file which allows for the addition of "trailers" for each correction. An analysis of the effect of the number of corrections on AGI discrepancy rates, and a study of the extent to which applicants had filed accurate AGIs, then changed them, were performed for the 1974-75 estimated income population.

Question 3.1: Does the frequency of corrections relate to the size of AGI discrepancies?

Cross-tabulations between both AGI and the reason for filing supplemental applications and the discrepancy in AGI were performed to examine this question.

As was true for the eligible population the vast majority of individuals has either one or two trailers. Discrepancy rates usually were similar for those submitting either one or two trailers, with discrepancies less frequent in high income categories for those with one trailer than for those with two. Discrepancies were more frequent for death cited as a reason if only one trailer was present, and less frequent for the unable to pursue income related activities if only one trailer was present.

Less than ten percent of the ineligible population had more than two trailers. Non-match rates due to data suppression were quite high and confounded analysis considerably. Therefore, it is not known if the categories of reason or income are related to the number of corrections.

Because the number of ineligible individuals who estimated their income and who filed three or more corrections was so small, it may be assumed that the number of corrections does not affect the program to any large extent. Since it is common for persons filing supplementals to do so as a correction to their original application for the year, the number of two trailer records was expected to be large, as it was for both the eligible and ineligible estimated income populations.

Question 3.2: Do any applicants report accurate AGIs, then submit corrections in order to increase the size of their awards?

In order to determine the answer to this question, tables were generated to show:

- How many individuals with more than one trailer who were currently discrepant had filed an accurate AGI on a previous trailer
- How many individuals were found to be ineligible based on a correction filed immediately before the supplemental

Both of these analyses yielded study groups smaller than 100 individuals; the small sizes and data suppression precluded any meaningful analysis. However, the fact that the numbers were so small indicates that the corrections process is rarely used by the ineligible estimated income group to increase the size of their award or to become eligible.

QUESTION #5: WHAT ARE THE EFFECTS OF MISREPORTING BY THE ESTIMATED INCOME POPULATION ON BEOG ELIGIBILITY AND THE SIZE OF AWARDS?

The background to this analysis was discussed in greater detail in section 4.3.2.2 Question #5. Briefly, this analysis was designed to examine the effect of discrepancies found in the investigation of other questions on the student eligibility index (SEI) of individuals in the population. The SEI indicates the size of student's award and hence, has important fiscal ramifications for the Basic Grant Program when applicants qualify for Grants

for which they would not qualify if they had utilized true representations of their financial characteristics.

For the estimated income population, discrepancies, even large ones, are expected, since applicants are essentially guessing about their future income for the year. The Validation Study was designed to determine the key factors of applicant characteristics which are related to discrepancies in order to improve accuracy of those applications using estimates of income. Many of the persons found to be ineligible for a Grant (SEI over 1200) may have qualified if they had been better able to estimate their true income.

By substituting the data applicants give to IRS for the data they report to BEOG on their supplemental application, we can analyze whether their SEI would have been affected if their year's true income had been used in the calculation of their SEI. Three data elements were substituted in series:

- Adjusted Gross Income
- Taxes Paid
- Household size

First, AGI was substituted. In a second analysis AGI and taxes paid were substituted. Finally, AGI, taxes paid and household size were substituted for the data applicants had provided to Basic Grants. Most of the discussions below will deal primarily with the AGI and AGI and taxes paid substitutions. Household size was definitionally quite different between the IRS and BEOG data bases. Therefore, these data were used only to provide added input to the results of other analyses.

Overall SEI Accuracy

Overall, ineligibles had greater SEI discrepancies, both positive and negative than did the eligibles, and a lower rate of

accuracy, given the insignificant discrepancy range of ± 200 SEI points. With 33.2 percent of the sample not matched, only 22 percent had insignificant discrepancies compared to 33 percent of the eligibles which had higher rates of non-match. Another 34.8 percent of ineligibles had positive discrepancies as opposed to the eligibles which had a positive discrepancy rate of 28 percent overall. In addition, the negative discrepancy rate for ineligibles was over nine percent, while eligibles had only three percent in this category. Exhibit 4.29 presents a summary of the SEI analysis for the ineligible population.

The distribution of individuals with SEI discrepancies differed by several percentage points from the distribution of AGI discrepancies for ineligibles, but there was a smaller difference between the SEI-AGI distributions than that encountered for the eligible population. The change in the ineligible population distribution is as follows:

Discrepancy Type	AGI Percentage	SEI Percentage (AGI Substitution)
Positive Discrepancy	65.5	52.1
Accurate (± 200)	14.5	32.9
Negative Discrepancy	20.0	14.9

By contrast, eligibles were slightly more discrepant on AGI than ineligibles, but considerably less discrepant for SEI than the ineligible group. For example, only 13.1 percent of eligibles had accurate AGIs, but 50 percent had accurate SEIs. This may be attributed to the nature of the Basic Grant formula. The formula is designed so that the lowest eligibility index is worth the largest award (assuming that cost of education and enrollment status warrant the maximum award). Since, by law, a Basic Grant cannot exceed \$1,400, the eligibility index has a lower limit - zero. Therefore,

EXHIBIT 4.29

SUMMARY OF OVERALL DISCREPANCY RATES IN
STUDENT ELIGIBILITY INDICES OBTAINED IN
IRS-SUBSTITUTION ANALYSES

SEI DISCREPANCY	SUBSTITUTION TYPE		
	AGI	AGI AND TAXES PAID	AGI TAXES PAID
POSITIVE DISCREPANCIES GREATER THAN +200	52.1%	53.6%	56.1%
ACCURATE (± 200)	32.9%	32.4%	29.9%
NEGATIVE DISCREPANCIES LESS THAN -200	14.9%	14.0%	13.9%

the formula sets negative values (caused by offsets greater than income and assets) to zero. Consequently, a portion of the eligibles have a negative index in absolute terms. Even if the IRS-reported income is larger than the BEOG-reported income, it can still be partially or completely absorbed by the offsets. This is probably the reason that eligibles are impacted less than ineligibles by income discrepancies.

For AGI substitutions, 22.6 percent of the ineligible population were in the over 800 point SEI discrepancy range. For AGI plus taxes paid 20.3 percent were in this range. Another interesting aspect of the AGI, taxes paid and household size combined substitution is that only an estimated 37 individuals had discrepancies in SEI of less than ± 1 , compared to a much higher frequency (1969) for eligibles. Indeed, the percentage and frequency of SEIs less than ± 1 was quite low for ineligibles as a whole: only 5.8 percent of the study group for the AGI substitution.

One of the explanations for the high degree of accuracy in SEI for the eligibles was that many of the incomes reported were low compared to ineligibles, resulting in little, if any, Federal Income Tax. Also, higher percentage of the ineligible population were dependents and had higher incomes.

Another interesting finding in the overall accuracy analysis is that negative discrepancies are not much more frequent for ineligibles than for eligibles. It would be expected that if ineligibles have a more difficult task in estimating (higher incomes, thus more room for error), that more should have overestimated. However, this was not the case. The frequency of negative discrepancies remained small. Thus, it would appear that most estimators tend to be conservative in their guesses, with the chance for errors, which would cause them to get smaller awards, minimized.

The ineligible population, like the eligibles, is less discrepant for SEI substitutions than it had been for the factors used for substitution. For the study group only 52.1 percent had positive discrepancies which were significant as compared to 64 percent for AGI. As previously explained, this is due in large part to the fact that insignificant AGIs had little effect on the SEI and therefore, should have resulted in substitution calculated-SEIs which were not as discrepant as the substitution variable. That is, one effect of using an insignificant SEI range is to expand the range of errors which are considered insignificant. If zero had been used as the criterion of accuracy, 27 percent more of the cases would have been considered discrepant, causing an increase in the discrepancy rate over the AGI analysis. However, the ± 200 point insignificant range was used to account for differences in variables and other factors affecting analysis. More detailed discussion of the substitution analysis was presented in section 4.3.2.2, Question #5.

As was the case for the eligibles, most of the discrepancies fell into the over 800-point range. However, there were larger frequencies of discrepancies in lower range categories than had been exhibited in analyses of other variables.

Dependency Status

Independents who were ineligible for Basic Grant awards were somewhat less discrepant than ineligible dependent applicants, in both under- and overestimating. A higher percentage of independents had accurate SEIs than did dependents.

Among the ineligibles there was, comparatively, a high rate of overestimation -- somewhat less than for the AGI analysis, but higher than that found for eligible applicants. Another difference between eligibles and ineligibles may be seen in the fact that, in the AGI substitution, over half of the eligible independents were accurate,

while only 32.9 percent of the ineligible independents were considered accurate. For the AGI and taxes paid substitutions, the percentage of accurate independents was 32.3 percent. See Exhibit 4.30.

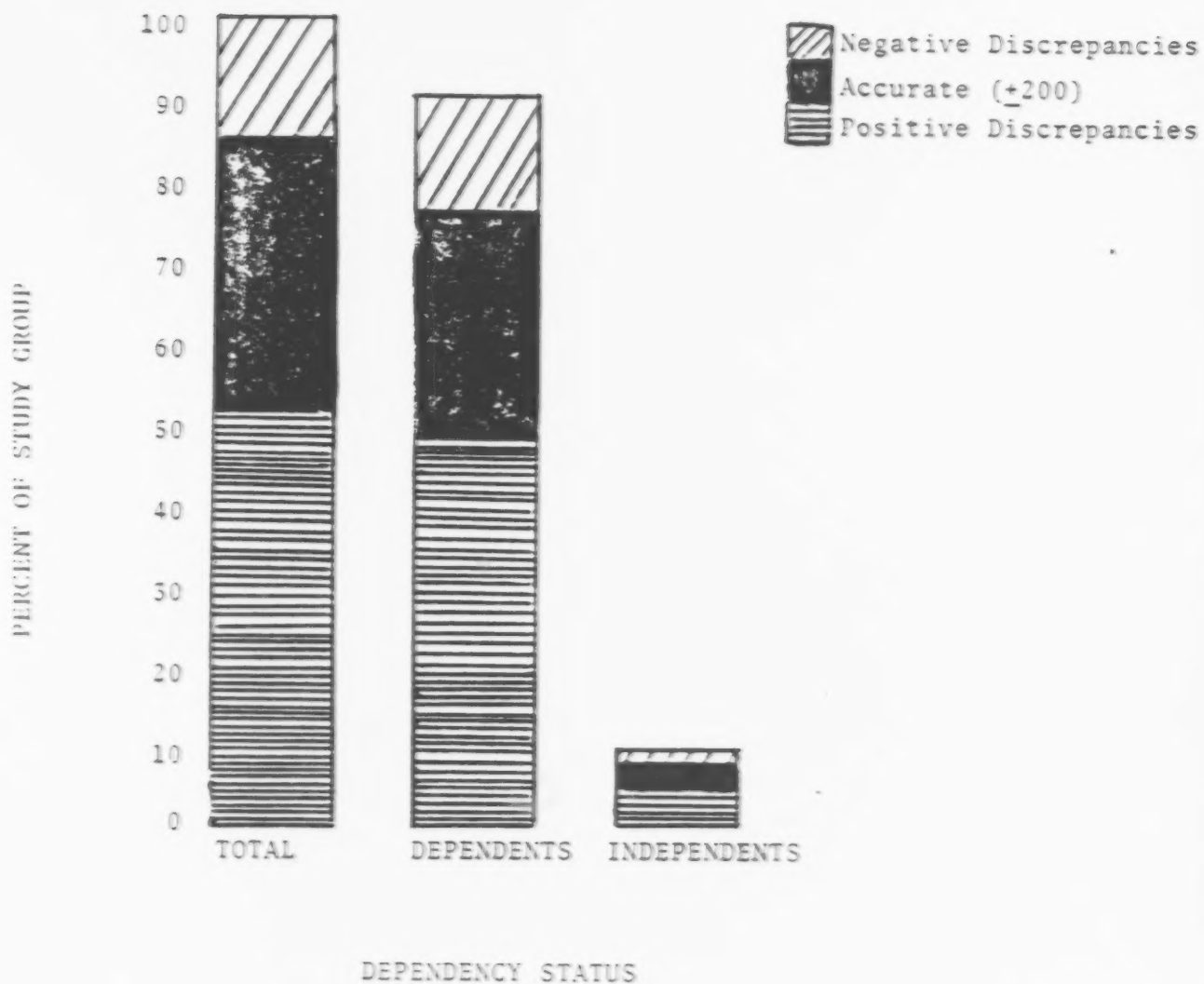
Adjusted Gross Income

Since it was known that dependents had higher adjusted gross income levels than independents separate analysis of the relationships of SEI discrepancies to adjusted gross income were performed. Exhibits 4.31 and 4.32 display the findings of AGI analysis.

The dependent applicants in the over \$12,500 AGI group accounted for 34 percent of the study group. Other income categories were much smaller, the largest being the under \$4,000 AGI group which had 15.8 percent of the study group. The AGI used in this analysis was the estimated figure supplied on the Basic Grant supplemental application. Several interesting findings should be highlighted:

- As income increased the percentage of positive discrepancies (resulting in greater eligibility) increased;
 - .. However, the differences in percentage of positive discrepancies over 800 points was not related to income level.
 - .. Small sized discrepancies generally increased in frequency as income levels increased.
- The rate of insignificant SEI discrepancies was not related to income level.
- The rate of negative discrepancies (resulting in lower eligibility) was strongly related to income. As income level increased, negative discrepancies increased in frequency. This is true for almost all levels of discrepancy.
- As income levels increased, the rate of matches in the population increased - from 52.4 percent for under \$4,000 income to 77.8 percent for over \$12,500 AGI.
- Across all income levels, the rate of accuracy was lower than it had been for eligibles.

EXHIBIT 4.30: FREQUENCY OF DISCREPANCIES IN STUDENT
ELIGIBILITY INDEX BY DEPENDENCY STATUS-
INELIGIBLES



4.94

EXHIBIT 4.31: FREQUENCY OF DISCREPANCIES IN STUDENT ELIGIBILITY INDICES BY ADJUSTED GROSS INCOME - DEPENDENT INELIGIBLES

SEI DISCREPANCY	BEOG-REPORTED AGI					
	TOTAL	UNDER \$4,000	\$4,000 to \$6,999	\$7,000 to \$9,999	\$10,000 to \$12,499	\$12,500 AND MORE
TOTAL STUDY GROUP	3912	618 15.8	597 15.3	729 18.6	636 16.3	1332 34.0
TOTAL POSITIVE DISCREPANCIES GREATER THAN +200	2078 53.1	399 10.2 64.6	312 8.0 52.3	357 9.1 48.9	346 8.8 54.4	664 16.9 49.8
TOTAL ACCURATE (+200)	1265 32.3	207 5.3 33.5	201 5.1 33.7	279 7.1 38.3	186 4.8 29.2	392 10.0 29.4
TOTAL NEGATIVE DISCREPANCIES LESS THAN -200	569 14.5	12 0.3 1.9	84 2.1 14.1	93 2.4 12.8	104 2.7 16.4	276 7.1 20.7

Count
Total Percent
Column Percent

EXHIBIT 4.32: FREQUENCY OF DISCREPANCIES IN STUDENT ELIGIBILITY INDICES BY ADJUSTED GROSS INCOME - INDEPENDENT INELIGIBLES

SEI DISCREPANCY	BEOG - REPORTED AGI				
	TOTAL	UNDER \$4,000	\$4,000 to \$6,999	\$7,000 to \$9,999	\$10,000 or MORE
TOTAL STUDY GROUP	349	137 39.2	111 31.8	40 11.5	12 3.4
TOTAL POSITIVE DISCREPANCIES GREATER THAN +200	197 56.4	73 20.9 53.3	61 17.5 54.9	14 40.0 35.0	*
TOTAL ACCURATE (+200)	92 26.3	53 15.2 38.7	28 8.0 25.2	11 3.2 27.5	*
TOTAL NEGATIVE DISCREPANCIES LESS THAN -200	60 17.2	11 3.2 8.0	22 6.3 19.8	15 4.3 37.5	12 3.4 100.0

Count
Total Percent
Column Percent

BEST COPY AVAILABLE

172

171

In the analysis of AGI discrepancies and income level in Questions 1 and 2, ineligible had not been substantially different in accuracy rates than eligibles, but, in terms of SEI discrepancies, they were significantly less accurate. The income levels as reported to Basic Grants are considerably higher for ineligible, with the over \$12,500 group accounting for 34 percent of ineligible, but less than 5 percent of the eligibles.

Interestingly, the over +800-point SEI discrepancy category was slightly smaller for the eligibles, although overall positive discrepancy rates were the same for both populations.

For independents accuracy rates decreased substantially for ineligible compared to eligibles. Ineligible independents had more individuals in high income categories than did eligibles. The rate of positive discrepancies did not appear to be related to income level in the independent ineligible population. Accuracy levels did decrease with increases in incomes and negative discrepancies increased with increases in income. Positive discrepancies, however were stable across income levels with the exception of the \$7,000 - \$9,999 group which had substantially fewer-positive discrepancies than any other group. However, the non-match rate of this group and the over \$10,000 income level group were so high (over 70 percent in the AGI and taxes paid substitution) that generalizations as to the accuracy of these groups would incorporate substantial error.

Student Eligibility Index

In order to determine the effect on Grant awards of discrepancies found through substitution of IRS data, it was of interest to compare discrepancy rates for the levels of SEI obtained by the ineligible population through the information they reported to Basic Grants on their supplemental applications. Exhibit 4.33 illustrates these results.

EXHIBIT 4.33: FREQUENCY OF DISCREPANCIES IN STUDENT ELIGIBILITY INDICES BY SEI
DISCREPANCY - INELIGIBLES

SEI DISCREPANCY RANGE	BEOG-REPORTED SEI					
	TOTAL	1,200 TO 1,400	1,401 TO 1,600	1,601 TO 1,800	1,801 TO 2,000	2,000 OR MORE
TOTAL STUDY GROUP	4,349	551 12.7	548 12.6	464 10.7	411 9.5	2,375 54.6
TOTAL POSITIVE DISCREPANCIES GREATER THAN +200	2,277 52.3	271 6.2 49.1	304 7.0 55.5	245 5.6 52.8	202 4.6 49.1	1,255 28.9 52.8
TOTAL ACCURATE (+200)	1,445 33.2	226 5.2 41.0	183 4.2 33.4	185 4.3 39.9	149 3.4 36.3	702 16.1 29.6
TOTAL NEGATIVE DISCREPANCIES LESS THAN -200	627 14.4	54 1.2 9.8	61 1.4 11.1	34 0.8 7.3	60 1.4 14.6	418 9.6 17.6

COUNT
TOTAL %
COLUMN %

Ineligibles were evenly divided among the lower eligibility index groups, but the 2000-or-more-point group accounted for 34.6 percent of the total study group. More than 52 percent of the ineligibles were found to have positive discrepancies. The frequency of positive discrepancies did not appear to be strongly related to the SEI level of the applicants. Insignificant discrepancies accounted for 33.2 percent of the study group. The 1200 to 1400 point group had the highest rate of accuracy with 41 percent in this discrepancy category, and the over 2000-point group had the lowest rate with 29.6 percent. Negative discrepancies accounted for 14.4 percent of the study group. Most of the negative discrepancies occurred in the two highest SEI groups (those over 1800 points) with the 1600 to 1800-point group having the lowest frequency. The AGI and taxes paid analysis had similar results.

IRS - Eligible Applicants

As has thus far been shown, the ineligible study group had a higher frequency of negative discrepancies in their SEIs than had the eligibles. This could mean that some ineligibles would have qualified to receive a Grant if they had been able to estimate their incomes more accurately. The number of individuals who would have received SEIs of less than 1200, and therefore become eligible, if IRS data had been substituted for their own estimates, was tabulated for each of the substitutions. A total of 290 individuals would have been eligible to receive a Grant if they had been able to estimate their AGIs correctly. This accounts for 6.6 percent of the ineligible study group. Of these, 82.4 percent were dependents and 17.6 percent were independents. As income level increased, the frequency of dependents who would have been eligible increased; the frequency of "new" eligibles decreased for AGI plus taxes paid and the AGI, taxes paid, and household size substitutions to 3.6 percent and 5.3 percent of the

study group. The characteristics of the individuals who would have qualified remained the same. (See Exhibit 4.34.) It might be remembered that almost 25 percent of the eligible study group would have lost their eligibility based on IRS data. Thus, the discrepancies obtained for the ineligibles have a very small impact on the Basic Grant Program. Only 5 or 6 percent of the ineligible study group was affected by their inaccuracies in terms of their ineligibility.

AGI Discrepancy Range

The final analyses performed for the eligibility index substitution were a series of cross-tabulations with the purpose of determining if the AGI discrepancy ranges obtained for applicants were in any way related to the SEI discrepancy ranges. Because the objective was to delineate those discrepancies which were related to one another, small ranges of AGI discrepancies were utilized. This had the unfortunate effect of preventing detailed analysis because of the extent of the suppression of data in cells displaying less than ten individuals, because the ineligible estimated income group was comparatively small.

In general, the number of positive SEI discrepancies increased as the AGI discrepancy ranges increased in size. Negative discrepancies could not be evaluated. The frequency of accurate SEIs was highest for the lowest AGI ranges, comprising more than half of the groups in the AGI ranges of discrepancies smaller than \$1000, and much smaller portions of larger discrepancy ranges.

Question #3: Summary

For ineligibles the differences in frequency between AGI discrepancies and SEI discrepancies were not as large as they had been for the eligibles study group. A comparatively small number of ineligibles would have been eligible for a grant if they had been better able to estimate. Thus, it would appear that, among the ineligibles, the estimation process did not significantly affect the results of their student eligibility index.

EXHIBIT 4.34: NUMBER AND PERCENT OF BEOG-REPORTED INELIGIBLE AND IRS-REPORTED ELIGIBLE APPLICANTS WITH ESTIMATED INCOME BY BEOG-REPORTED AGI AND DEPENDENCY STATUS

DEPENDENCY STATUS	BEOG-REPORTED AGI					
	TOTAL	LESS THAN \$4,000	\$4,000 TO \$6,999	\$7,000 TO \$9,999	\$10,000 TO \$12,499	\$12,500 TO \$ MORE
TOTAL STUDY GROUP	237 100.0	29 12.2	46 19.4	46 19.4	48 20.3	68 28.7
DEPENDENT	COUNT 197 ROW PERCENT 100.0 COLUMN PERCENT 83.1 TOTAL PERCENT	15 7.6 51.7 6.3	33 16.8 71.8 13.9	33 16.8 71.8 13.9	48 24.4 100.0 20.3	68 34.5 100.0 28.7
INDEPENDENT	47 100.0 19.8 19.2	14 29.8 48.3 5.9	13 27.7 28.2 5.5	13 27.7 28.2 5.5	*	*

NOTE: INDEPENDENT APPLICANTS WITH BEOG-REPORTED AGI EXCEEDING \$10,000 ARE REPORTED IN THE \$10,000 TO \$12,499 STRATUM.

NOTE: ELIGIBILITY INDICES REPORTED IN THIS TABLE ARE BASED UPON SUBSTITUTION OF THE IRS-REPORTED AGI, AND TAXES PAID FOR THE CORRESPONDING BEOG-REPORTED DATA.

4.4: SUMMARY AND CONCLUSIONS

The Estimated Income IRS Sample component of the Basic Grant Validation Study utilized the best available criterion of accuracy in its assessment of the accuracy of estimations made on certain Basic Grant applications. This criterion of accuracy was the Internal Revenue Service data base. By matching a sample of BEOG estimated income applicants with corresponding data on the IRS master file, a broad investigation of the estimated income process could be undertaken. This investigation was limited to overall study of applicant characteristics and was prevented from more detailed analysis by the need to maintain the strict privacy of individuals IRS and BEOG data records.

The IRS data regarding income are considered the most valid indicator of family financial characteristics. For the purposes of the present study these data were assumed to be accurate. Other components of the Estimated Income study used less valid criteria. For example, the CRT Look-up Study compared the 1974-75 estimate to the 1975-76 report of income. The 1975-76 BEOG data were not necessarily accurate--no validation of them was made and the error rate could have been substantial. Indeed, as a part of the Individual Case Follow-up a study of 1975-76 data available for 1974-75 estimators found these data to be discrepant in more than half of the cases where the 1974-75 estimates were also discrepant.

The IRS sample study utilized a stratified random sampling technique to select BEOG records. The high rate of non-matches between IRS and BEOG data, however, could not be accounted for since individuals in the study could not be identified personally. Non-match was approximately 35 percent for the total population. This portion was assumed to be randomly distributed, but was not included in the analysis. Tables were produced to present the matched "study group only and were inflated so that the sampling proportions were equalized.

A high rate of errors, particularly underestimates were found for both eligible and ineligible applicants. The ineligibles were slightly more accurate than the eligibles, and reflected higher income levels. Ineligibles also include a higher proportion of applicants citing death of a parent or spouse as the reason why they filed applications.

Errors in estimating adjusted gross income were made by 86.4 percent of the total study group (86.9 percent for eligibles and 86 percent for ineligibles). Most of the discrepancies were positive indicating that 68.5 percent of the total study group had estimated that they would have a smaller AGI for the year than they actually had. Eligibles had a rate of underestimation of 70.3 percent; ineligibles had a 64 percent rate of underestimation. Overestimation accounted for 17.8 percent of the total study groups; eligibles overestimated 16.6 percent of the time; ineligibles, 20 percent.

Adjusted gross income discrepancies were related slightly to income, with the sizes of discrepancies increasing as income increased. Independents were more accurate than dependents, but not by large amounts. The supplemental reason categories also were related to discrepancies, but not by large differences in discrepancy rates. Those citing unable to pursue income activities were the most accurate; those citing death were the least accurate, especially for ineligibles. Only the date of signature for the applicants' latest change to their 1974-75 BEOG record was strongly related to the accuracy of their estimates. Accuracy increased sharply for the last two date categories, November 1 to December 31 and after December 31.

Discrepancies in taxes paid occurred much less frequently than AGI discrepancies, and discrepancies in household size/exempting occurred for less than ten percent of the study group.

Discrepancies formed in any of the above analyses, however, do not directly influence the student eligibility indices for the applicants. Adjusted gross income taxes paid and household size are all elements in the SEI calculations, but none has a "one for one" effect on the level of SEI. Therefore, each of these data elements as they existed in the IRS master file was substituted in the SEI formula for the corresponding BEOG reported elements, and the indices were recalculated. The SEI determines the eligibility of the applicant and the maximum size of his/her award. The effects of the IRS substitution were then evaluated.

The major finding of the SEI analysis was that approximately half of the eligible applicants studied had insignificant discrepancies in their SEIs, and approximately one-quarter had no discrepancy in SEI at all based on the IRS substitution. For the ineligible study group, it was found that accuracy of SEI was increased with the substitution but not as much as for eligibles. However, an analysis of whether the IRS substitution would change ineligible applicants from eligible to ineligible or vice versa, showed that only 5 percent of the ineligibles would have been eligible if they had been better able to estimate. Thus, the estimation process is exceptionally accurate for ineligibles. At the same time 25 percent of the eligibles studied would have been disqualified on the basis of IRS data. Three factors are significant when drawing conclusions about the extent of errors found in the estimation process. First, applicants were limited in their ability to estimate their expected future income. Second, although the rates of discrepancies in AGI were quite high, the effects of these discrepancies on the eligibility index were substantially lower than the rate of AGI discrepancies would have suggested. And, third, the estimated income population as a whole represented less than two percent of all BEOG applicants; estimated income applicants with discrepancies accounted for about one percent of all BEOG applicants.

In addition to the investigation of discrepancies and their impact on the student eligibility index, several other questions were studied using the IRS Sample data. The results of these are discussed below:

- DOES THE FREQUENCY OF CORRECTIONS MADE BY APPLICANTS TO THEIR BEOG APPLICATIONS RELATE IN ANY WAY TO MIS-REPORTING?

Among the estimated income population over 85 percent of BEOG records had either one or two trailers. Because the supplemental is often filed as a correction to the original application, and because the group having second trailer was more accurate than those with only one, it was concluded that the second trailers of the estimated income population represented their supplemental applications. The remainder of the study group had a higher degree of errors, but a very high non-match rate in comparison to the others. Those with more than three trailers were not numerous enough for analysis. Also, there was not a significant enough number of cases for which there was a correction made and a previously recorded accurate AGI in order to make an assessment as to whether this technique was utilized for increasing BEOG eligibility.

- DO APPLICANTS SUPPLY TRUE COPIES OF THE FORMS WHEN PARTICIPANTS IN THE VALIDATION EFFORTS?

Because of the large numbers of individuals participating in the Validation effort who had discrepancies between the tax forms they provided and their BEOG applications it was concluded that Validation participants for the Estimated Income Individual Case Follow-up components did, in fact, supply true copies of their tax forms. This was corroborated by the IRS Sample study which showed a small difference in the size of discrepancies found in the two components studies. This difference was approximately 7 percent.

and because of the known error of .05 percent at the .95 confidence interval for the sample obtained for IRS, and the unknown error rates for the Individual Case Follow-up and the IRS non-matches the difference was considered to be insignificant.

Chapter Five which follows this discussion, will present a complete summary and the integration of findings from the three component studies.

5

OVERALL FINDINGS OF THE ESTIMATED INCOME STUDIES

The three Estimated Income Study components were designed to examine the accuracy of estimates made by those Basic Grant applicants who met the criteria necessary for filing supplemental applications in the 1974-75 academic year. Each of the studies used a different perspective in analyzing the accuracy of estimations and examined a different set of analysis questions.

Each of the three studies examined a separate set of analysis questions. Two of the questions were investigated by each component. The common analysis questions were:

- What is the overall accuracy of estimation?
- What is the relationship between the applicant's characteristics and the accuracy of estimation?

Other analyses were common to only two of the studies, because of the range of data items available to each. Where possible, analysis questions were addressed in more than one of the studies in order to cross-verify the findings through analysis using different samples and slightly different populations.

Summary of Findings

The major findings of the three component studies may be highlighted as follows:

- The accuracy of estimation was strongly associated with the latest date of signature of applicants' records. The later the date, the more accurate the estimate.
- Estimating accuracy was less strongly related to the level of income (as estimated), dependency status, the reason for filing the supplemental, and eligibility.
 - .. accuracy decreased slightly with increases in estimated income
 - .. independents were slightly more accurate than dependents
 - .. ineligibles were more accurate than eligibles
 - .. persons citing "unable to pursue income activities" as their supplemental reason tended to be more accurate than other estimators
 - .. the larger the "true" income, the greater the incidence of underestimating; the smaller the "true" income, the greater the incidence of overestimating.^{1/}

Statistically, the distribution of application AGI discrepancies in the three studies may be summarized as follows:

	CRT Look-up	Individual Validation	IRS Study.*
Estimated income and Actual Income within:			
± \$200	24.7%	-	13.1%
± \$300	28.5%	17.0%	16.4%
± \$1,000	47.3%	36.3%	37.1%
± \$2,000	64.7%	-	57.6%
± \$3,000	-	67.5%	-

* Eligibles only to allow comparison between studies.

^{1/} This finding is consistent with findings of Boyd and Himelick, and Orwig and Jones in studies of the accuracy of income estimates for populations (unlike the Basic Grant estimating population) comprised largely of persons who did not suffer any substantial decrease in income in the current year.

Variations in Rates of Accuracy and Study Designs

As is shown in the above exhibit, all of the studies found that three-fourths of the applicants were found to have discrepancies in their income estimates. However, each study found a different rate of accuracy. The variations in rates of accuracy may be attributed to two factors in addition to sampling variations - one, each study utilized a criterion for accuracy with a different degree of reliability; and two, each study had limitations in the design that may have caused some distortions in the findings.

The criteria for determining the accuracy of the income estimates used in each of the three studies follow:

Study	Criterion for Accuracy
CRT Look-up Study	Actual 1974 adjusted gross income reported by applicant on regular 1975-76 application
Individual Case Follow-up Study	Adjusted gross income on a copy of 1974 Federal Income Tax Return supplied by applicant or parent
IRS Sample Study	Applicant's or parents' 1974 adjusted gross income reported to and maintained by IRS

Comparisons between the study findings showed that the criterion used in the CRT Study - 1974 income reported on 1975 application - was less reliable than the criteria used in the other two studies. The reliability of the CRT criterion was studied in the Individual Case Follow-up. A sample of cases with estimating errors were examined to determine whether the income reported on the 1975 application was accurate according to the tax form supplied by the applicant. It was found that discrepancies existed for three-fourths of these applicants and that about half of the discrepancies were greater than \$200. Not only did this finding indicate that the actual 1974 income reported on the 1975-76 application was the least reliable

of all criteria for accuracy in this study, but it also showed that the applicants had supplied true copies of tax returns. Otherwise, few discrepancies would have been found between the tax return and the 1975-76 application. The IRS Study provided further testimony to the validity of the tax forms supplied by applicants since the findings of the IRS Study and the Individual Case Follow-up were quite similar.

In addition to the differences in criteria for accuracy used in the three studies, unique features of each study design may have caused inconsistencies in findings. In the CRT study, only applicants who had filed an income estimate in 1974-75 and a regular application in 1975-76 were sampled for study. In the IRS and Individual Case studies, 1974-75 estimators were sampled, regardless of whether they filed at all in 1975-76. The inequality of these populations (and samples) may have accounted for some of the variation in the findings.

The Individual Validation Study was unique in that it afforded direct contact with the applicants and parents. Therefore, it was possible to obtain clarification about any questionable information supplied. For example, it was found that some applicants included both taxable and nontaxable income in their AGI estimate. In other cases, a joint tax form was supplied from divorced parents and the income of only one parent was required for the estimate; and the AGI estimate may have included the incomes of married parents filing separate returns. Additional clarifying information for these types of circumstances enhanced the validity of Individual Validation findings. They were not available, however, for the CRT and IRS Study.

The IRS Study entailed computerized analyses of the AGI reported by the independent applicant or the dependent applicant's parent whose social security number was reported on the BEOG application.

Only one IRS record was matched and it was therefore impossible to determine cases of divorced parents filing a joint return and cases where nontaxable income had been included in the estimate.

The nonresponse/non-match rate may also impact on the validity of the findings. It was found that the non-matched records in the CRT Study and the nonrespondents in the Individual Case Follow-up tended not to differ from the study groups. Consequently, it was concluded that the error associated with these factors was minimized and distributed randomly. In the IRS Study, however, matches were successful for about half the records attempted, and the matched group did differ from the total sample in some respects. The matched study group contained relatively fewer independents and fewer low income applicants than the population. It was expected that this occurred because low income applicants (and many independents who also report low incomes) did not match because their income were too low to necessitate filing a tax return. This may have resulted in a bias in the IRS Study since the low income estimators who did not file tax returns probably estimated fairly accurately.

In conclusion, it appears that the Individual Case Follow-up Study and the IRS Study better estimate accuracy of reporting for the BEOG population than the CRT Study. Since the IRS and Individual Follow-up Studies obtained fairly consistent findings, they provide the best indicators of the accuracy of income estimates reported to Basic Grants.

Impact of Discrepancies on Eligibility

While the extent of discrepancies in AGI appear to be extremely large, AGI did not have a direct impact on the eligibility index of the applicants. Income is the major variable on which the index is based, but other variables such as taxes paid and household size are also considered. In fact, taxes paid and household size

were found to be substantially more accurate than AGI for all categories of applicants studied. To examine the composite impact of reporting and estimating discrepancies of these variables on the eligibility index, the IRS Sample Study substituted IRS versions of these three variables (AGI, taxes paid and household size) for the BEOG reported data. The results of the substitution analysis on eligible applicants found that less than half of those who underestimated their income had significant discrepancies in their eligibility. Indeed, the substitution of both AGI and taxes paid resulted in only 43.6 percent of the applicants having significant changes (± 200 points) in eligibility indices. There was also a substantial drop in significant overestimators. For ineligibles it was found that only 3 percent of the estimators would have been eligible based on substituted data.

Overall, the estimation process incorporates much higher rates of inaccuracy than exists for the regular application process. However, the extent of errors in eligibility indices is smaller compared to errors found for AGI. The estimated income population represents a special subgroup of BEOG applicants with rather unique characteristics regarding their financial circumstances. In addition, the total number of estimators amounts to less than two percent of the BEOG population. Therefore, it may be concluded that errors found in this group did not have a significant impact on the Program. Because estimators have suffered traumatic circumstances affecting their abilities to pay for their postsecondary education, it is recommended that no changes be made to the nature of the estimation process in the BEOG program.